REPORT RESUMES

ED 002 980 24

THE EFFECTIVENESS OF SPECIAL DAY CLASS TRAINING PROGRAMS FOR SEVERELY (TRAINABLE) MENTALLY RETARDED CHILDREN.

BY- DUNN, LLOYD M. HOTTEL, JOHN V.
GEORGE PEABODY COLL. FOR TEACHERS, NASHVILLE, TENN.
REPORT NUMBER CRP-103 PUB DATE 58
CONTRACT OEC-SAE-6425
EDRS PRICE MF-$0.18 HC-$3.32 83P.

DESCRIPTORS- *SPECIAL CLASSES, MENTAL RETARDATION, *RETARDED CHILDREN, *BEHAVIOR DEVELOPMENT, *VOCABULARY DEVELOPMENT, *INTELLECTUAL DEVELOPMENT, PEABODY PICTURE VOCABULARY TEST, NASHVILLE, TENNESSEE

THE OBJECTIVE WAS TO TEST TWO HYPOTHESES--(1) DAY CLASS TRAINING WOULD PRODUCE DESIRABLE CHANGES IN THE BEHAVIOR OF THE TRAINABLE MENTALLY RETARDED CHILDREN OVER AND ABOVE THOSE WHICH RESULT WHEN THE CHILD REMAINS AT HOME AND HAS NO TRAINING IN A SCHOOL SETTING, AND (2) SCHOOL PLACEMENT OF SUCH CHILDREN WOULD PRODUCE DESIRABLE CHANGES IN THE HOME ADJUSTMENT OF THE FAMILIES INVOLVED. DURING PHASE 1 A PICTURE VOCABULARY SCALE WAS DEVELOPED TO MEASURE CHANGES IN THIS ASPECT OF THE DEVELOPMENT OF TRAINABLE CHILDREN. THEN THE EXPERIMENTAL DESIGN OF THE STUDY WAS PERFECTED, SUBJECTS WERE IDENTIFIED, AND A PILOT STUDY WAS MADE CENTERING ON THE TWO HYPOTHESES. PHASE 2 OF THE STUDY CARRIED OUT THE EXPERIMENTAL DESIGN. IN SPITE OF SOME POSITIVE TRENDS, THERE WAS AN INDICATION THAT THE CHILDREN IN THE SPECIAL CLASSES DID NOT MAKE MORE PROGRESS IN SOCIALIZATION, INTELLECTUAL DEVELOPMENT, SELF-CARE, OR PARENT-CHILD RELATIONS THAN THE CHILDREN WHO REMAINED AT HOME. HOWEVER, THE SPECIAL CLASSES AS OPERATED APPEARED TO PRODUCE GREATER CHANGES IN THE HIGH IQ GROUP. LITTLE EVIDENCE ACCRUED TO SUPPORT THE CONTENTION THAT DAY CLASS TRAINING FOR TRAINABLE CHILDREN, AS PRESENTLY CONSTITUTED, IS EFFECTIVE. (GC)

THE EFFECTIVENESS OF SPECIAL DAY CLASS TRAINING PROGRAMS FOR SEVERELY (TRAINABLE) MENTALLY RETARDED CHILDREN

January 1, 1957-June 30, 1958

Lloyd M. Dunn, Project Director

PEABODY COLLEGE RESEARCH SERIES IN MENTAL RETARDATION

GEORGE PEABODY COLLEGE FOR TEACHERS

NASHVILLE 5, TENNESSEE

THE EFFECTIVENESS OF SPECIAL DAY CLASS TRAINING PROGRAMS FOR SEVERELY (TRAINABLE) MENTALLY RETARDED CHILDREN*

Lloyd M. Dunn, Project Director

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

SECTION I

Report of a Study of the Effectiveness of Special Day Classes for Trainable Children

John V. Hottel, Principal Investigator

SECTION II

Development of the "Peabody Picture Vocabulary Test"

Lloyd M. Dunn, Principal Investigator

January 1, 1957 - June 30, 1958

ED 002980

*The research reported herein was performed pursuant to Contract SAE-6425, United States Office of Education, Department of Health, Education, and Welfare

1958

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education

3414a
75

SECTION I

Report of a Study of the Effectiveness of Special Day Classes for Trainable Children

John V. Hottel, Principal Investigator

ACKNOWLEDGEMENTS

Recognition is extended the Research Advisory Committee who provided continuous consultative service (singly and as a group) to the project staff. This committee was comprised of the following personnel from the Tennessee State Department of Education: Ruth B. Anderson, Supervisor, Program for Trainable Children; Ronald E. Brinkley, Director of Special Education; James E. Hogan, Supervisor of Special Education; and William E. Turner, Coordinator of the Division of Instruction. The following faculty of Peabody College also served on the committee: Gordon N. Cantor, Associate Professor of Experimental Psychology; Rue L. Cromwell, Associate Professor of Clinical Psychology; Margaret Hudson, Assistant Professor of Special Education; Raymond C. Norris, Associate Professor of Psychology; and Wallace A. Verburg, Associate Professor of Special Education.

Appreciation is gratefully expressed to Alfred L. Baldwin, Head of the Department of Child Development and Family Relationships, Cornell University; The Fels Research Institute for the Study of Human Development; and the American Psychological Association for permission to use the modified version of the Fels Scales of Parent Behavior.

Appreciation is also expressed to Dr. Samuel A. Kirk, Director, Institute for Research on Exceptional Children, University of Illinois for permission to revise The Illinois Behavior Rating Scale, and for reading the interpretation of data.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS. iii

LIST OF TABLES AND FIGURES. v

FOREWORD. vii

Chapter

I EXPERIMENTAL PROCEDURES AND RESULTS 1

- Hypotheses 1
- Subjects 1
  - Comparisons of the Groups. 2
  - The Experimental Group 3
  - The Control Group. 4
  - Matching of Pairs. 4
- Instruments. 4
  - Scoring Procedure. 6
- Procedure. 7
- Findings of the Study. 9
- Comparisons of Total Experimental and Control Groups. . . . 9
  - Basic Measures 9
  - Behavior Rating Scale. 9
  - Modified Fels Scales of Parent Behavior. 10
- Comparisons of *Post Hoc* Sub-Groups 15
  - Mental Age 15
  - Intelligence Quotients 17
  - Social Ages. 18
  - Social Quotients 19

II SUMMARY, DISCUSSION, AND CONCLUSIONS. 20

- Hypotheses 20
- Findings 20
- Discussion 21
- Conclusions. 24

LIST OF TABLES AND FIGURES

Table		Page
1	Distribution of the Matched Pairs by Sex and Clinical Types	3
2	Initial Comparisons of the Experimental and Control Groups on Basic Measures	5
3	Comparisons of the Total Experimental and Control Groups on Basic Measure Mean Change-Scores (1957 Score Minus 1956 Score).	10
4	Comparison of Total Experimental and Control Groups on Behavior Rating Scales Mean Change-Scores (1957 Score Minus 1956 Score).	12
5	Comparison of Total Experimental and Control Groups on Fels Scale Mean Change-Scores (1957 Score Minus 1956 Score).	14
6	Comparisons of the Sub-Groups' Mean Change-Scores Binet Mental Ages (in Months) (1957 Scores Minus 1956 Scores)	16
7	Summary of Analysis of Variance on Mean Change-Scores Binet Mental Ages (in Months) (1957 Scores Minus 1956 Scores)	16
8	Comparisons of the Sub-Groups' Mean Change-Scores Binet Intelligence Quotients (1957 Scores Minus 1956 Scores)	17
9	Summary of Analysis of Variance on Mean Change-Scores Binet Intelligence Quotients	17
10	Comparisons of the Sub-Groups' Mean Change Scores Vineland Social Age (1957 Scores Minus 1956 Scores). .	18
11	Summary of Analysis of Variance on Mean Change-Scores Vineland Social Age (1957 Scores Minus 1956 Scores). .	18

Table Page

12 Comparisons of the Sub-Groups' Mean Change-Scores Vineland Social Quotients (1957 Score Minus 1956 Score) 19

13 Summary of Analysis of Variance on Mean Change-Scores Vineland Social Quotients (1957 Scores Minus 1956 Scores). 19

Figure

1 Comparisons of the Total Experimental and Control Groups on Basic Measure Mean Change-Scores (1957 Scores Minus 1956 Scores) 10

2 Comparison of Total Experimental and Control Groups on Behavior Rating Scales Mean Change-Scores (1957 Scores Minus 1956 Scores). 11

3 Comparisons of the Total Experimental and Control Groups on Modified Fels Scales Mean Change-Scores (1957 Scores Minus 1956 Scores). 13

Foreword

In recent years children with severely retarded intellectual development (I.Q. range 30 to 50) have posed a serious challenge to a number of public agencies--especially our public schools. Prior to 1950 these boys and girls were either placed in state residential schools or were retained in the homes of their parents. They were denied a public day school education because they had been found to be uneducable academically when placed in a regular classroom or in a special class for the educable mentally handicapped. Parents and families of these children frequently were ostracized by society. As a reaction to these conditions, in 1950, parents and friends of the retarded began banding together. They actively campaigned for greater social awareness and acceptance of these children. One of their most frequent requests was for the establishment of special day classes for trainable children--with a curriculum geared especially to their needs and capacities. The requests of these parents' groups precipitated within many communities a shifting of responsibility among agencies--public health, mental health, social welfare, and public instruction--largely because scientific evidence was not available to indicate which public agency or profession was best prepared to serve these children and their families.

In spite of lack of needed research information, parent demands increased to such a point that many states, including Tennessee, passed legislation providing state financial support for special day classes for trainable children. This rapidly expanding (and controversial) movement has been questioned by many educators. Some believe the progress of these children in self-care and social adjustment will not be accelerated by such placement. There was a critical need for a controlled study of the effectiveness of special day school programs, as compared with home training programs, to provide objective evidence as to the best placement for these youngsters.

George Peabody College for Teachers entered into agreements with the Tennessee State Department of Education and the U.S. Office of Education to conduct such a study. The study was primarily financed by grants from the State of Tennessee for the period July 1, 1955 through December 31, 1956 and by funds from the Federal Government for the period January 1, 1957 through June 30, 1958.

During the school year 1955-1956, 23 classes for "trainable" children were established within 16 school systems throughout the State of Tennessee, pilot and descriptive studies were conducted, and experimental and control groups established. (For a detailed description of the first year of operation, the composition of the classes, and the descriptive study, see the Interim Report.[a]) The experimental study was carried out during the 1956-1957 school year by which time 19 school systems had established programs.

---

[a] John V. Hottel. The Tennessee Experimental Program of Day-classes for Severely Mentally Retarded (Trainable) Children: Interim Report of the Study. Nashville: George Peabody College for Teachers. 1956.

CHAPTER I

EXPERIMENTAL PROCEDURES AND RESULTS

Hypotheses

Three assumptions were tested by the experimental research reported here. The two ad hoc hypotheses were as follows:

a. Day class training would produce desirable changes in the behavior of the children themselves;

b. Day class training would produce desirable changes in parent-child relations in the homes of the children involved.

It was decided by the study staff that the program could be considered effective if either or both of these positions were substantiated. (Of course it was recognized that the ultimate effectiveness of the special classes would rest on how well children from such classes adjusted to home and community in their adolescence and adulthood.)

A third post hoc hypothesis grew out of an analysis of the data collected in testing the ad hoc positions. This was as follows:

c. Children with I.Q.'s 41 to 50 in day class training would make significantly more progress than other sub-groups of children studied.

Subjects

The children who were to become beneficiaries of the day class program were to be "severely mentally retarded". This term was defined in Public Chapter No. 33 as follows:[1]

> For the purpose of the Act, the term "severely mentally retarded children" shall mean children who have attained the chronological age of six years but who have not reached their twenty-first birthday and who, because of mental retardation, are not considered to be of educable mind but who may be expected to benefit from training in group settings designed to further their social adjustment and economic usefulness in their own homes or in a sheltered environment and who are capable of communicating satisfactorily and of participating in groups.

This broad definition was amplified by the Rules, Regulations, and Minimum Standards of the Tennessee State Board of Education, where admission to the State's day class program was limited to children who: (a) possessed the ability to respond to simple directions, (b) possessed the ability to communicate his needs, (c) was not considered dangerous to himself or to others, and (d) was capable of participating in groups. These criteria constituted the framework within which State Department personnel based their recommendations for placement of children in the special classes. However, additional stipulations were set forth limiting eligibility of children for inclusion into the experimental aspects of the study. In addition to meeting the general criteria listed above, children in the experimental group must be attending special class for the first time beginning with the 1956-57 school year, and children included in either the "experimental" or "control" groups must: (a) possess a Binet intelligence quotient between 30 and 50, and (b) be between the ages of six and 16 years.

---

[1] Hereafter in this report the terms "severely mentally retarded" and "trainable" will be used interchangeably.

Only 31 children enrolled for the 1956-57 school year met the criteria listed above. They were selected as potential members of the "experimental" group. Forty children who remained at home (and who met all the criteria listed above) were identified as possible "controls" from approximately 200 families contacted during the summer months.

It was originally planned that 30 pairs of children would be matched on seven variables and then studied. (For every child attending a special class for the first time, a child would be selected who did not have access to day class training.) Matching was to be on the basis of: (a) sex, (b) clinical type or etiology of defect, (c) chronological age, (d) mental age, (e) social age, (f) socio-economic status of family, and (g) general cultural area. The restricted enrollment of children in the day class program for the 1956-57 school year made it necessary to modify the research design. It was possible to secure only 21 pairs of children matched on all but two ("f" and "g") of these preselected variables.[2] (Hereafter in this report children who were enrolled in the day classes will be designated the "experimental group" and the matching children who remained at home will be designated the "control group.")

Comparisons of the Groups: Comparisons of the children on sex and clinical types are presented in tabular form in Table 1. Identical matches were secured on the sex variable: eight pairs of the children (38 per cent) were female and thirteen pairs (62 per cent) were male. By grouping etiologies into broad categories--i.e., all types of organic brain injuries into an inclusive category here labeled "brain injury (undifferentiated)"--identical matches were also obtained on this variable. Twelve pairs of children (57 per cent) were diagnosed as brain injury (undifferentiated); six pairs (29 per cent) as mongoloid; and three pairs (14 per cent) as of unknown etiology.

---

[2] The 21 matched pairs of children described in this report differ considerably from the 20 matched pairs presented in the Interim Report. (See footnote a.) Rematching was made necessary by changes in the structure of the original groups. Several children from each of the original matched pairs moved out of the state during the school year; others were never entered into the classes or were dropped soon after entrance. Some of the members of the control group were entered into private day-schools and thus became ineligible for the study. Others of the control group moved into areas where public day-classes were available. Children from the experimental group who were never entered

TABLE I

DISTRIBUTION OF THE MATCHED PAIRS BY SEX AND CLINICAL TYPES

Classification	Number	Per Cent
Female	8	38.0
Male	13	62.0
Total	21	100.0
Brain injury (undifferentiated)	12	57.1
Mongoloid	6	28.6
Unknown etiology	3	14.3

The following basic measures were obtained for each member of the experimental and control groups: chronological age, Binet mental age and intelligence quotient, and Vineland social age and social quotient. These data are presented in Table 2.

The Experimental Group: As shown in Table 2, the chronological ages (C.A.'s) of the children in the experimental group ranged from 70 to 187 months (5.8 to 15.5 years) with a mean of 103.9 months (8.7 years). The Binet mental ages (M.A.'s) of the group ranged from 24 to 71 months (2.0 to 5.9 years) with a mean of 42.4 months (3.5 years). The Binet intelligence quotients (I.Q.'s) of the experimental group ranged from 30 to 50 with a mean of 41.1. The range of Vineland social ages (S.A.'s) was from 28 to 107 months, with a mean of 57.9 months (4.8 years). Vineland social quotients (S.Q.'s) for the group ranged from 37 to 84, with a mean of 56.7.

---

in the classes were transferred to the control group. Children from the control group who entered public day-classes were transferred to the experimental group. Rematching was then accomplished on the basis of the original criterion measures secured for all the children, and the 21 matched pairs who are the subjects of this report resulted. Comparison of these 21 pairs with the 20 pairs reported in the Interim Report shows that the final 21 pairs are more closely matched than were the original pairs. Comparative statistics were not run on the composite groups reported in the Interim Report because drop-outs had reduced the total number of children in these groups to only slightly above those in the matched groups.

The Control Group: It can be seen in Table 2 that the C.A.'s of the children in the control group ranged from 68 to 192 months (5.7 to 16.0 years) with a mean of 103.0 months (8.6 years). M.A.'s within the group ranged from 24 to 86 months (2.0 to 7.2 years) with a mean of 41.4 months (3.4 years). I.Q.'s ranged from 30 to 49 with a mean of 39.9. S.A.'s of the control group ranged from 27 to 150 months, with a mean of 54.3 months (4.5 years). S.Q.'s ranged from 35 to 98 with a mean of 50.1 (See Table 2).

Matching of Pairs: The experimental and control groups were matched on each of the basic measures described above. This was accomplished by means of t tests for correlated measures (after Helen M. Walker and Joseph Lev. Statistical Inference. New York: Henry Hold, 1953, p. 156). As shown in Table 2, in no case did the t-ratio approach the critical value of t at the .05 level of significance. (Throughout this study this level of significance was maintained.)

It should be noted that the 21 children of the experimental group represented only seven of the 19 school systems which had established classes for trainable children. No children who met the pre-established criteria were enrolled in classes of the other 12 systems at the beginning of the 1956-57 school year. The classes represented in the experimental study were distributed as follows: three in West Tennessee contributed four children to the experimental group; two in Middle Tennessee contributed eight children; and two in East Tennessee contributed nine children.

The 21 children of the control group were selected from 13 communities throughout the State. Seven of these were located in West Tennessee; eleven children of the control group lived in this area. Three of the communities were located in Middle Tennessee; four children came from there. The three remaining centers were located in East Tennessee; and six children came from that area.

Instruments

The Revised Stanford-Binet Intelligence Scales: Form L. 1937 Revision[3] and the Vineland Social Maturity Scale[4] were used to obtain the basic measures described in the preceding section. The results from these two scales were used both as criterion measures for matching the groups and as indicators of development of behavioral skills for the children. In addition to these two scales, a Rating Scale of Child Behavior,[5] which was intended to measure twelve dimensions of behavioral

---

[3] Lewis M. Terman and Maud A. Merrill. Measuring Intelligence. New York: Houghton Mifflin, 1937.

[4] Edgar A. Doll. Measurement of Social Competence. Nashville: Educational Test Bureau, 1953.

[5] This scale was developed by the research staff at Peabody College to measure growth in the areas of training specified in the Rules, Regulations, and Minimum Standards of the State Board of Education.

TABLE 2

INITIAL COMPARISONS OF THE EXPERIMENTAL AND CONTROL GROUPS ON BASIC MEASURES

Pairs (N= 21)	Code	Sex	Clinical Type	Chronological age*			Mental age*			Intelligence quotient			Social age*			Social quotient		
				Cont.	Exper.	Diff.	Cont.	Exper.	Diff.	Cont.	Exper.	Diff.	Cont.	Exper.	Diff.	Cont.	Exper.	Diff.
A	1	F	Brain injury	80	100	-20	33	43	-10	41	43	- 2	79	45	34	98	45	53
B	3	F	Brain injury	91	96	- 5	27	29	- 2	30	30	0	45	40	5	49	42	7
C	5	F	Brain injury	86	92	- 6	35	44	- 9	41	48	- 7	35	51	-16	41	55	-14
D	6	F	Brain injury	92	90	2	34	29	5	37	32	5	38	55	-17	41	61	-20
E	7	M	Brain injury	68	74	- 6	26	27	- 1	39	36	3	40	42	- 2	59	57	2
F	10	M	Brain injury	110	95	15	51	47	4	41	50	- 4	50	52	- 2	45	55	-10
G	13	M	Brain injury	101	102	- 1	43	44	- 1	43	43	0	64	60	4	63	58	5
H	14	M	Brain injury	113	109	4	50	51	- 1	44	47	- 3	51	83	-32	45	76	-31
I	16	M	Brain injury	121	125	- 4	37	47	-10	31	38	- 7	39	71	-32	35	58	-23
J	18	M	Brain injury	192	187	5	86	71	15	48	40	8	150	107	43	77	57	20
K	19	F	Brain injury	113	103	10	49	44	5	43	43	0	60	47	13	53	46	7
L	20	F	Brain injury	128	139	-11	54	50	4	42	36	6	59	72	-13	46	52	- 6
M	4	F	Mongolism	85	91	- 6	31	45	-14	37	49	-12	48	68	-20	57	75	-18
N	9	M	Mongolism	74	74	0	24	24	0	32	32	0	29	28	1	39	38	1
O	11	M	Mongolism	85	83	2	27	37	-10	32	45	-13	41	50	- 9	48	60	-12
P	12	M	Mongolism	89	103	-14	36	34	2	40	33	7	39	38	1	44	37	7
Q	17	M	Mongolism	115	108	7	50	50	0	43	46	- 3	36	46	-10	40	43	- 3
R	21	M	Mongolism	71	70	1	28	31	- 3	39	44	- 5	27	56	-29	38	79	-41
S	2	F	Unknown	166	150	16	73	64	9	45	43	2	119	58	61	73	39	34
T	8	M	Unknown	75	78	- 3	37	35	2	49	45	4	39	65	-26	53	84	-31
U	15	M	Unknown	108	113	- 5	39	44	- 5	36	39	- 3	53	82	-29	49	73	-24
Mean				103.0	103.9	-0.90	41.4	42.4	-1.00	39.9	41.1	-1.2	54.3	57.9	-3.6	52.1	56.7	-4.6
Standard deviation				30.9	27.9		15.7	11.7		5.5	6.0		29.7	18.1		15.2	14.1	
_t_ - ratio						0.47			0.62			1.03			0.66			0.94

*In months.

Critical value of _t_ at the .05 level of significance is 2.09.

skills, and a modified version of the Fels Scales of Parent Behavior,[6] which was intended to measure thirteen dimensions of parent-child relations and home adjustment, were used. (Copies of the Behavior Rating Scale and Fels Scales may be obtained from the project director).

Scoring Procedure: For the Binet, Vineland, and Fels scales, the scores reported in the "Results" section of this report were obtained by subtracting the 1956 measure from the 1957 measure, thus obtaining a difference score. The scoring procedure for the Rating Scale of Child Behavior was somewhat different. This scale purported to measure changes in the development of behavioral skills on a six-point, discrete interval scale going from dependent to independent behavior. The 1957 measure was compared with the 1956 measure and changes were reported on the individual sub-scales toward independent (plus) or toward dependent (minus) behavior. In this fashion, difference scores of as much as plus or minus five points were obtained for each item on this scale. Since this was an unstandardized instrument, no separate scores were reported for the initial and terminal evaluations.

On the Binet, Vineland and Behavior Rating Scales, a positive score represented a change toward desirable (independent) behavior, and a negative score represented a change toward undesirable (dependent) behavior. The modified Fels Scales could not be interpreted in the same manner. Ratings on each individual scale in this battery were accomplished by marking an appropriate point on a 90-millimeter vertical line. The score for each scale was the linear millimeter measure from the base of the line to the point of rating. "Good" adjustment was not always represented by the same point on the scale: on some items "good" adjustment was at the top of the scale, on others at the bottom, and on still others at the mid-point. Therefore, the positive and negative change-scores could not be interpreted per se as representing "desirable" and "undesirable" adjustment. It was necessary to interpret each of the sub-scales independently from the others. This precluded obtaining a "total adjustment score" by adding together the individual sub-scores, as was done for the Behavior Rating Scale. The modified Fels Scales consisted of thirteen individual sub-scales selected on the basis of highest inter-correlation with other scales in the battery but lowest inter-correlation with each other. The thirteen scales thus selected were:

---

[6] Alfred L. Baldwin, et al. "The appraisal of parent behavior." Psychol. Monogr. 63, 4, 1949, pp. 85.

Scale Number	Title
1.5	Discord in the Home
1.6a	Sociability of Family
1.7	Coordination of Household
1.91	Child-centeredness of Home
2.11	Duration of Contact
3.11	Restrictiveness of Regulations
3.12	Readiness of Enforcement
3.21	Quantity of Suggestion
3.3	Accelerational Attempt
5.1	Readiness of Criticism
5.2	Direction of Criticism
7.1	Solicitousness of Child's Welfare
8.1	Understanding of Child

The Behavior Rating Scale yielded sub-scores on the following dimensions of developmental skills:

Scale Number	Title
1.1	Work Habits
1.2	Oral Language Development
1.3	Written Language Development
1.4	Number Development
1.5	Art Development
1.6	Music Development
1.7	Physical Development
2.1	Toilet Training
2.2	Cleanliness
2.3	Eating Habits
3.1	Dressing
4.0	Social Interaction

The change-scores for the sub-scales were added together and a total behavior score obtained. In this manner 13 scores were obtained from this instrument.

The Binet and Vineland scales each yielded two scores: MA and IQ from the Binet, and SA and SQ from the Vineland. Thus measures were obtained on 30 dimensions of behavior and/or home adjustment.

Procedure

Broad objectives of the training program were set forth in Public Chapter No. 33 as follows: ". . . training in group settings designed to further their social adjustment and economic usefulness in their own homes or in a sheltered environment. . ." The broad objectives thus stated in the Act were made explicit in the Rules, Regulations, and Minimum Standards of the Tennessee State Board of Education. Training was to include activities in the following areas: (1) personal routines, (2) health routines, (3) care of property, (4) safety routines, (5) play performance, (6) social adjustment, (7) language development. It was realized that the foregoing categories were not mutually exclusive and that much overlap would occur in materials, methods, and activities. However, such an explicit statement of objectives provided a framework for comparable training of children and facilitated research concerning the effectiveness of the program.

Implementation of common objectives and comparable procedures was attempted in two ways. First, two two-day in-service training sessions were conducted for the 23 teachers of the special day classes during the 1955-56 school year. At these meetings common problems were discussed, and the teachers developed a *Curriculum Guidelines* using the areas of training stipulated by the Tennessee State Board of Education as a framework and their own experiences as guides. Two additional two-day sessions were held for these teachers by the State Department of Education during the 1956-57 school year. At these meetings the *Guidelines* were revised. Again the general areas of training previously described provided the framework, and the additional year of experiences gleaned by the teachers dictated the specific revisions. Thus each of the teachers possessed identical *Curriculum Guides*: each had participated in the development of the Guides, and it was assumed that each had comparable understandings concerning the aims and methods of training severely mentally retarded children.

The second attempt at statewide implementation of common procedures was the efforts of the Supervisor of the Program for Severely Mentally Retarded Children employed by The State Department of Education. The Supervisor periodically visited each of the classes for the express purpose of aiding the teachers in developing programs which were comparable to one another.

Children for both the experimental and control groups were identified and tested on the *Binet* and *Vineland* measures within 30 days of the beginning of the 1956-57 school year. In addition to these measures a psychological examiner visited for approximately two hours in the home of each child. During the course of this visit, the mother of the child, in direct response to questions from the examiner, supplied information concerning specific skills possessed by the child at that time. The examiner scored each behavorial item on the *Rating Scale of Child Behavior* on the basis of the information thus obtained. The examiner also rated the home on the modified *Fels Scales of Parent Behavior* at the time of this initial visit.

The children of the experimental group were subsequently enrolled in the public day classes and provided with approximately 175 days of school experiences. The children of the control group were not enrolled in classes but remained in their homes.

There was no further contact between the children and parents involved in the study and members of the research staff until the end of the 1956-57 school year. At this time, all measures were repeated for each child and each home in both the experimental and control groups.

Findings of the Study

The matched pair design of this study was selected so that the t-tests of correlated measures (as have been described previously) could be used in comparing the experimental and control groups. During the course of this analysis, it became apparent that the 21 matched pairs were distributed according to I.Q. in such a fashion as to render the data amenable to further, more detailed study employing analysis of variance procedures. Consequently, post hoc hypotheses of differences between groups based on level of intelligence were formulated and tested. The findings of the study will be presented in separate sections. First to be dealt with will be the findings of the comparisons of the total experimental and control groups by t tests of correlated measures. This will be followed by a presentation of the findings based on the more detailed analyses of the post hoc sub-groups.

Comparisons of Total Experimental and Control Groups

The t-test comparisons of the total experimental and control groups indicated that no significant differences existed between these groups on 29 of the 30 dimensions of behavior and adjustment under consideration. In other words, day-class training produced no differences in the rate of development of MA, IQ, SA, SQ, Behavioral Skills, or Parent-child Relations over and above that developed by home training. The one dimension on which a significant difference was found was concerned with the amount of time spent in direct contact between mother and retarded child. This difference had been anticipated and the measure was included partially because it could serve as a validity check on the other scales in the modified Fels battery.

Basic Measures: The mean change-scores on these measures are presented graphically in Figure 1. No significant differences on MA, IQ, SA or SQ were found between the total experimental and control groups.[7] The mean change-scores and the t ratios associated with each are presented in Table 3.

Behavior Rating Scale: The mean change-scores for the groups on the Behavior Rating Scale are presented graphically in Figure 2. No significant differences were found between the total experimental and control groups on any of the 12 subscales of this measure, or on the total score obtained by adding together the sub-scale scores. The mean change-scores, and the t ratios associated with each are presented in Table 4.

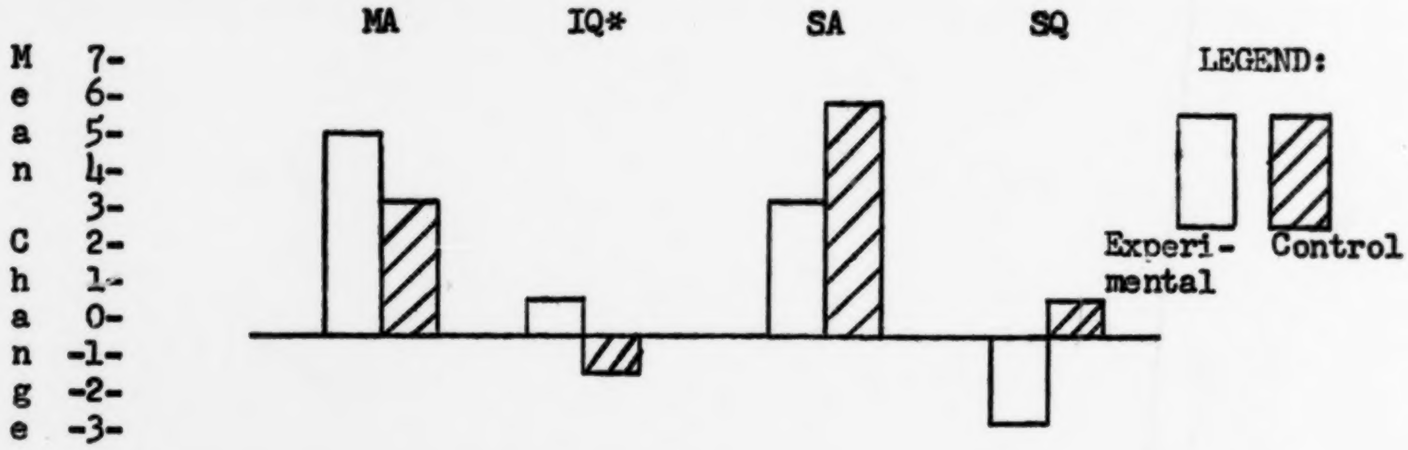
---

[7] Chronological age was held constant in this study; i.e., all initial measures were computed as of June 1, 1956 and terminal measures as of June 1, 1957. Therefore, IQ and SQ changes here become simple functions of MA and SA changes.

Modified Fels Scales of Parent Behavior: The mean change-scores for the Modified Fels Scales are presented graphically in Figure 3, and the comparisons of the scores with the associated t-ratios are given in Table 5. As has been stated, only one Fels sub-scale yielded a significant difference between the change-scores of the total experimental and control groups. It was the measure of the amount of time spent in direct contact between parent and retarded child--Fels Scale 2.11, "Duration of Contact".

Figure 1

COMPARISONS OF THE TOTAL EXPERIMENTAL AND CONTROL GROUPS
ON BASIC MEASURE MEAN CHANGE-SCORES (1957 SCORES MINUS 1956 SCORES)



* Losses in IQ and SQ associated with gains in MA and SA are explained by their ratio function between MA and CA, and SA and CA. IQ and SQ scores are measures of rate of change. Where the rate of change in MA or SA was not sufficient to maintain the initial ratio between these measures and rate of CA change, the associated IQ and SQ change-scores reflected losses.

TABLE 3

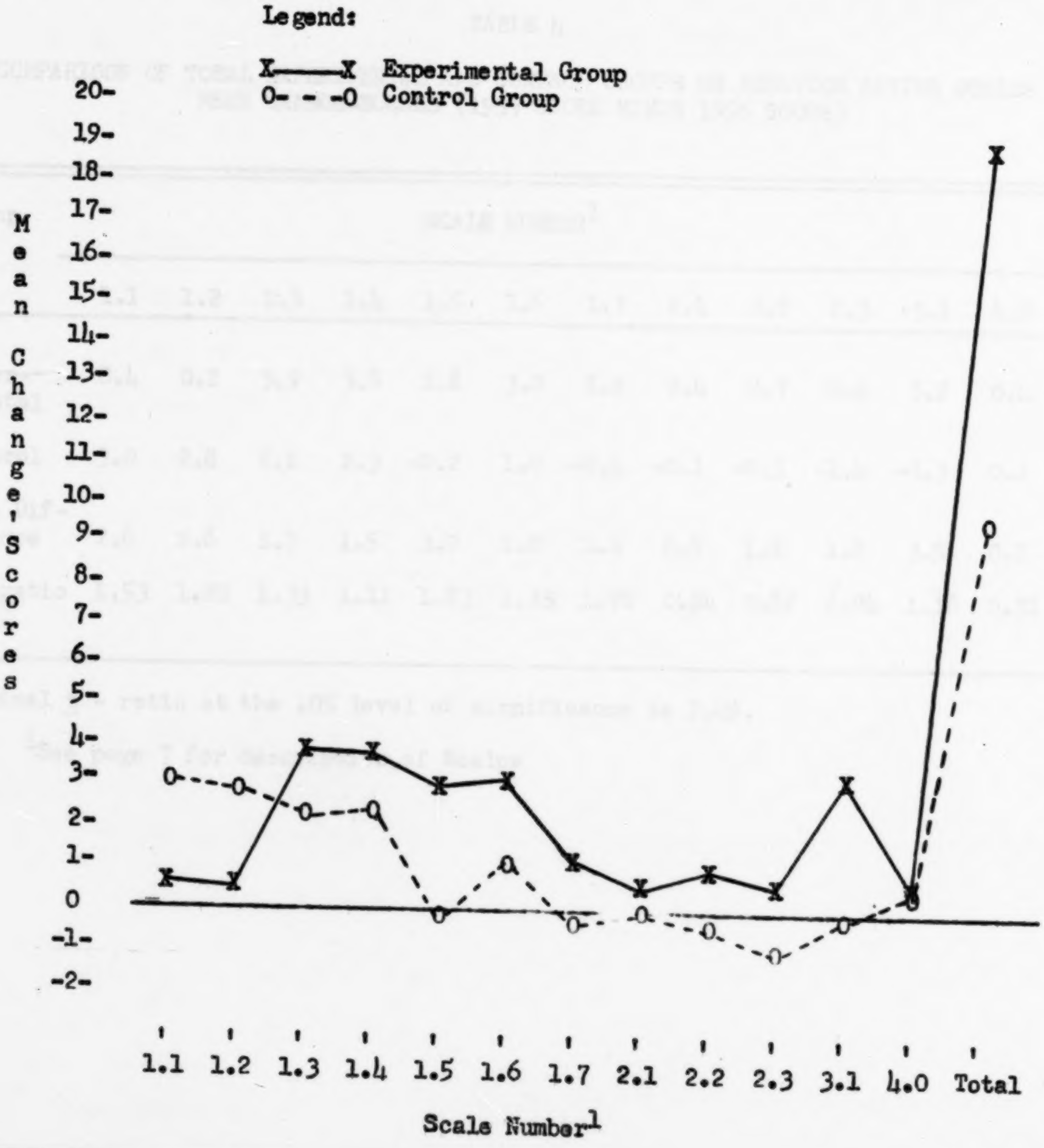
COMPARISONS OF THE TOTAL EXPERIMENTAL AND CONTROL GROUPS
ON BASIC MEASURE MEAN CHANGE-SCORES (1957 SCORES MINUS 1956 SCORES)

Group	Mental[1] Age	Intelligence Quotient	Social[1] Age	Social Quotient
Experimental	5.50	0.48	3.95	-2.81
Control	3.30	-0.62	6.67	0.33
Mean Difference	2.20	1.10	2.72	3.14
t ratio	0.66	0.75	0.75	0.95

[1] Change scores are expressed in months for MA and SA. Critical value of t at the .05 level of significance is 2.09.

Figure 2

COMPARISON OF TOTAL EXPERIMENTAL AND CONTROL GROUPS ON BEHAVIOR RATING SCALES MEAN CHANGE-SCORES (1957 SCORES MINUS 1956 SCORES)



---

[1] See page 7 for description of scales.

TABLE 4

COMPARISON OF TOTAL EXPERIMENTAL AND CONTROL GROUPS ON BEHAVIOR RATING SCALES MEAN CHANGE-SCORES (1957 SCORE MINUS 1956 SCORE)

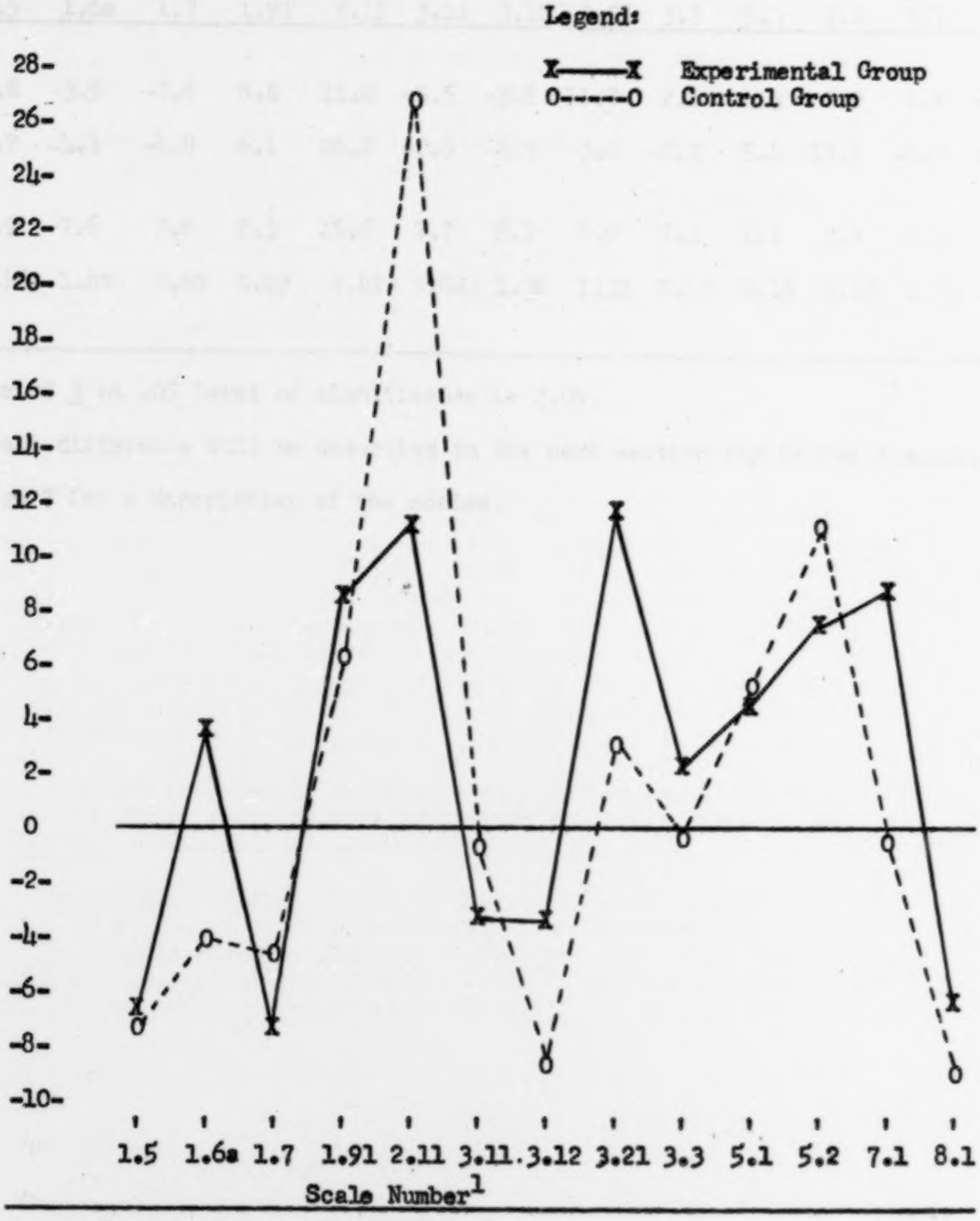
Group	SCALE NUMBER[1]												
	1.1	1.2	1.3	1.4	1.5	1.6	1.7	2.1	2.2	2.3	3.1	4.0	Total
Exper-imental	0.4	0.2	3.9	3.8	2.8	3.0	1.2	0.4	0.7	0.4	3.2	0.4	19.0
Control	3.0	2.8	2.2	2.3	-0.2	1.0	-0.4	-0.1	-0.5	-1.4	-0.3	0.2	9.4
Mean Dif-ference	2.6	2.6	1.7	1.5	3.0	2.0	1.6	0.5	1.2	1.8	3.5	0.2	9.6
_t_ - ratio	1.53	1.20	1.33	1.11	1.83	1.33	1.96	0.84	0.82	2.04	1.18	0.21	1.17

Critical _t_ - ratio at the .05 level of significance is 2.09.

[1] See page 7 for description of Scales

Figure 3

COMPARISONS OF THE TOTAL EXPERIMENTAL AND CONTROL GROUPS ON MODIFIED FELS SCALES MEAN CHANGE-SCORES (1957 SCORES MINUS 1956 SCORES)



---

[1] See page 7 for description of scales.

TABLE 5

COMPARISON OF TOTAL EXPERIMENTAL AND CONTROL GROUPS ON FELS SCALE MEAN CHANGE-SCORES (1957 SCORE MINUS 1956 SCORE)

Group	SCALE NUMBER[1]												
	1.5	1.6a	1.7	1.91	2.11	3.11	3.12	3.21	3.3	5.1	5.2	7.1	8.1
Experimental	-6.8	-3.5	-7.6	8.4	11.0	-3.5	-3.8	11.7	2.1	4.3	7.9	8.9	-6.5
Control	-7.7	-4.1	-4.8	6.1	26.6	-0.8	-8.7	3.0	-0.2	5.4	11.1	-0.5	-9.0
Mean Difference	0.9	7.6	2.8	2.3	15.6	2.7	5.1	8.0	2.3	1.1	3.2	9.4	2.5
t - ratio	0.46	1.82	0.62	0.49	2.42*	0.003	1.38	1.32	0.50	0.18	1.15	1.71	0.35

Critical value of t at .05 level of significance is 2.09.

*This significant difference will be described in the next section and in the discussion.

[1] See page 7 for a description of the scales.

Comparisons of Post Hoc Sub-Groups

During the course of the original analysis, it was observed that experimental children who were higher in intelligence appeared to be exhibiting gains whereas children with lower IQ's were not. In other words, there appeared to be a differential in gains in relation to differences in intellectual ability. To test this post hoc hypothesis, the experimental and control subjects were divided into two sub-groups: 11 pairs with IQ's between 30 and 40 (hereafter referred to as the "low" IQ group) and 10 pairs with IQ's between 41 and 50 (hereafter referred to as the "high" IQ group).

A Type I analysis of variance (after E.F. Lindquist. Design and Analysis of Experiments in Psychology and Education. Boston: Houghton-Mifflin, 1953, pp. 266-267.) was performed for each of the four basic measures which made up the thirty dimensions described in the preceding sections. This type of analysis was used because the "matched-pairs" design of the experiment made it possible to treat the data as if they represented two scores obtained from different treatments for the same individual. The analysis made it possible to compare the groups and sub-groups in the following manner: (a) the main effects between the experimental and control total groups; (b) the main effects between the total high and total low IQ groups; and (c) the interaction effects between type of treatment--i.e., day class training versus home training--and level of intelligence. A significant interaction on a dimension rendered the main effects for that measure meaningless. In such cases, it was necessary to perform t-tests of the simple effects to determine the source of interaction. (Because t-test comparisons of the total experimental and control groups had already been performed and reported in the previous section, main effects between the two groups will not be included in this section.) The matched-pairs design necessitated the use of two types of t-tests: a test of correlated measures when comparing the experimental and control groups (described previously); and a test of non-correlated measures when comparing the high and low IQ sub-groups (after Walker and Lev. Op. cit. p. 157).

Although analyses of variance were computed for all four measures used in the study, only the results concerning the analyses of the Binet and Vineland scales will be discussed here.

Mental Age: A significant interaction (level of intelligence interacting with type of treatment) was found for mental age scores: therefore t-tests were performed to determine the source. The mean mental age change-scores are shown in Table 6, and the summary of the analysis of variance is presented in Table 7.

TABLE 6

COMPARISONS OF THE SUB-GROUPS' MEAN CHANGE-SCORES BINET MENTAL AGES (IN MONTHS) (1957 SCORES MINUS 1956 SCORES)

Group	Experimental	Control	Mean Difference
High IQ	8.5	0.6	7.9
Low IQ	2.8	5.1	2.3

TABLE 7

SUMMARY OF ANALYSIS OF VARIANCE ON MEAN CHANGE-SCORES BINET MENTAL AGES (IN MONTHS) (1957 SCORES MINUS 1956 SCORES)

Source	Variance	F
Experimental vs Control	69.43	2.68
High vs Low IQ	3.71	0.19
Interaction	271.03	10.48*

*$p < .05$ Critical value of F at .05 level of significance is 4.38.

The mean change-scores presented in Table 6 were compared by means of t-tests. Differences between the sub-groups were as follows: (1) children with IQ's between 41 and 50 made significantly greater gains in Binet mental age when they attended day classes than when they remained at home without formal training; (2) children with IQ's between 41 and 50 made significantly greater gains in Binet mental age than children with IQ's between 30 and 40 when both groups attended special classes; (3) children with IQ's between 41 and 50 did not progress any more rapidly in mental age than children with IQ's between 30 and 40 when both groups remained at home; and (4) children with IQ's between 30 and 40 who attended day classes progressed no more in mental age than children in the same IQ range who remained at home.

Intelligence Quotients: Chronological ages for all children in this study advanced exactly one year; therefore, IQ is a simple function of MA. Since changes in IQ bear exactly the same relationship to one another as do changes in MA, the results outlined above for MA apply to IQ. The t-test comparisons of the mean IQ change-scores and the summary of the analysis of variance for IQ's are presented in Tables 8 and 9 respectively.

TABLE 8

COMPARISONS OF THE SUB-GROUPS' MEAN CHANGE-SCORES
BINET INTELLIGENCE QUOTIENTS (1957 SCORES MINUS 1956 SCORES)

Group	Experimental	Control	Mean Difference
High IQ	2.9	-2.3	5.2
Low IQ	-1.7	0.9	2.6

TABLE 9

SUMMARY OF ANALYSIS OF VARIANCE ON MEAN CHANGE-SCORES
BINET INTELLIGENCE QUOTIENTS

Source	Variance	F
Experimental vs Control	12.60	0.83
High vs Low IQ	5.27	0.33
Interaction	160.83	10.57*

*$p < .05$ Critical value of F at .05 level of significance is 4.32.

Social Ages: No significant differences were found between either the total high and low IQ groups or the group sub-divisions. Mean social age change-scores are presented in Table 10, and the summary of the analysis of variance in Table 11.

TABLE 10

COMPARISONS OF THE SUB-GROUPS' MEAN CHANGE-SCORES
VINELAND SOCIAL AGE (1957 SCORES MINUS 1956 SCORES)

Group	Experimental	Control	Total
High IQ	8.5	8.4	8.4
Low IQ	-0.2	5.1	2.5
Total	4.0	6.7	---

TABLE 11

SUMMARY OF ANALYSIS OF VARIANCE ON MEAN CHANGE-SCORES
VINELAND SOCIAL AGE (1957 SCORES MINUS 1956 SCORES)

Source	Variance	F
Experimental vs Control	77.36	0.56
High vs Low IQ	376.57	1.92
Interaction	75.60	0.54

Critical value of F at the .05 level of significance is 4.38.

Social Quotients: Social quotients were a simple function of social age in this study because chronological age was held constant. Like intelligence quotients, they are not discussed separately. The mean change-scores for social quotients are presented in Table 12, and the summary of the analysis of variance is presented in Table 13.

TABLE 12

COMPARISONS OF THE SUB-GROUPS' MEAN CHANGE-SCORES
VINELAND SOCIAL QUOTIENTS (1957 SCORE MINUS 1956 SCORE)

Group	Experimental	Control	Total
High IQ	1.2	1.1	1.2
Low IQ	-5.6	-0.4	-2.9
Total	-2.3	0.3	---

TABLE 13

SUMMARY OF ANALYSIS OF VARIANCE ON MEAN CHANGE-SCORES
VINELAND SOCIAL QUOTIENTS (1957 SCORES MINUS 1956 SCORES)

Source	Variance	F
Experimental vs Control	74.67	0.64
High vs Low IQ	176.50	0.97
Interaction	73.06	0.63

Critical value of F at the .05 level of significance is 4.38.

CHAPTER II

SUMMARY, DISCUSSION, AND CONCLUSIONS

Under contracts with the Tennessee State Department of Education and the U.S. Office of Education, George Peabody College for Teachers undertook to conduct a two-year study of the effectiveness of special day classes for "trainable" children which were supported by funds appropriated by the 1955 Tennessee State Legislature.

During the first school year (1955-56), 23 classes for trainable children were established in 16 school systems throughout the state, serving 245 pupils. In this first year, pilot and descriptive studies were conducted, and experimental and control groups were selected in readiness for an intensive study the following year.

During the second year (1956-57), a controlled study was made of 21 pairs of children, one member of each pair being a child enrolled in a special class for the first time at the beginning of the 1956-57 school year and the other member being a comparable child not enrolled in a special class. The purpose of this study was to compare the progress of the "school" (or experimental) group with the "home" (or control) group. The social, intellectual, and self-care skills as well as parent-child adjustment of these 21 pairs of children were evaluated by psychological and educational tests prior to the beginning of the 1957-58 school year, and again just before the end of this school year.

Hypotheses: Two ad hoc hypotheses were tested: (a) day class training would produce significant and desirable changes in the children themselves, and/or (b) would produce significant and desirable changes in parent behavior toward the children. (It was agreed by the study staff that day classes would be considered effective if either, or both, of these hypotheses were substantiated.)

A post hoc hypothesis grew out of analysis of the data: (c) children with I.Q.'s 41 to 50 in day class training would make significantly more progress than other sub-groups of children studied.

Findings: A statistical comparison was made between all 21 children of the "school" group, and all 21 children of the "home" group to test the two ad hoc hypotheses. In social, intellectual, and self-care skills, it was found that the "school" group did not make gains large enough during the year to be statistically significant from the changes made by the "home" group. Similarly, changes in parent-child relations made by the two groups were not statistically significant. In other words, neither of the original hypotheses were substantiated.

Examination of the performances of the individual children in the "school" group seemed to indicate that some had made substantial gains, but these gains were obscured by the performance of other children who showed little progress or even retrogressed slightly during the school year. In general children with higher I.Q.'s appeared to be making the most progress. To explore this more fully a statistical comparison was made between sub-groups of both the "school" and "home" groups. The 21 pairs of children were divided into four groups, yielding a 2 x 2 analysis of variance design. Both the "school" and "home" subjects were divided into two sub-groups: 11 pairs with I.Q.'s between 30 and 40 (referred to as the "low" I.Q. group) and 10 pairs with I.Q.'s between 41 and 50 (referred to as the "high" I.Q. group). When these groups were statistically compared, it was found that children with the higher I.Q.'s who attended special classes gained significantly more in mental age than either children with comparable intelligence who remained at home or children with lower intelligence who attended the special classes. Conversely, children with the lower I.Q.'s who attended school did not make gains in mental growth greater than children of comparable intellectual levels who remained at home. In other words the post hoc hypothesis was partially substantiated in that special classes promoted the mental growth of children with an I.Q. over 41 but did not promote the mental growth of children with a lower I.Q. This increase in mental age for the high I.Q. group who attended special day classes may well be the most important finding of the study.

Discussion: The findings of this study were influenced by both limiting and strengthening factors. Some of the limitations were as follows:

1. The number of children in the study (only 21 pairs of pupils) was small. Significant results may have been found with larger groups. In addition, the small sample made it impossible to compare the effectiveness of the various teachers in the program, or to group the children according to the type of mental retardation involved, e.g., brain-injured children may have changed in a different fashion to mongoloid children.

2. The psychological and educational tests available to the study staff were, in general, not sensitive to small changes in the behavior of the children or in parent-child relations. Perhaps the most sensitive instrument was the Revised Stanford-Binet Intelligence Scale. The Modified Fels Scale of Parent Behavior and the Vineland Social Maturity Scale appeared to measure only gross changes. The Behavior Rating Scale was especially weak. On the one hand, the nature of the scale made it impossible for children who were rated near the independent behavior end of the scale on the initial rating to exhibit increments (in other words, the scale did not have sufficient top). On the other hand, children who were rated near the dependent-behavior end of the scale on the initial rating had ample opportunity to move upward (toward independent behavior) on the terminal rating. Further, it was the opinion of the study staff that instruments used in the study measured inadequately the results of some major aspects of training, particularly socialization, which was emphasized in some of the classes. We cannot be assured that major changes did not occur just because the tests were unable to identify and measure them.

3. Another limitation of the study was the probability of interrater differences. Ideally the initial and final measures on the subjects

should have been made by one examiner. Because of the practical problems of geographical distances and the tight time schedule, it was necessary to employ seven psychological examiners to collect the initial measures on possible subjects for the experimental group. In a period of approximately three and one-half months, almost 100 applicants for admission to the special classes for the first time, effective Fall 1956, were examined and their homes visited. Three of the examiners collected data on a pool of children (approximately 200) for the control group. The terminal (retest) data were collected by two of the psychologists involved in the initial testing plus an additional examiner who was not involved the first year. This resulted in a total of eight psychological examiners contributing to the inter-rater variance. Because of the small number of subjects, it was impossible to analyze the data to determine the effect of inter-rater differences.

4. The length of time involved in the control study may have been too short to allow changes to take place. The gains made by the children in as short a time as one school year (actually only about 175 school days) may have been overshadowed by the insensitiveness of the measuring tools. This would suggest that future studies should extend over a longer period of time. (It was not possible to continue the present study because a number of the "home" group was enrolled in special classes at the beginning of the 1958 school year.)

5. Another limiting factor must be considered in interpreting the results. The treatment (or training) of the experimental subjects may not have been repjresentative of children, in general, in the special day classes of the State of Tennessee. No objective data are available to refute or substantiate this possibility; however, subjective judgments were obtained. Observations were made of each class by the principal investigator to determine the composition of the classes, curriculum, and training techniques employed by the various teachers. Several conclusions were drawn from these observations: (1) The numerical enrollments of the classes and the range of CA's and MA's in each class precluded the assumption of comparability of classes on these dimensions. (2) The physical location of classrooms and the facilities and materials available to the children were different for each class. (3) It was noted that teachers were not utilizing comparable materials or training techniques in all the classes. (4) The general orientations of some of the classes were academically oriented with major emphasis on reading and writing skills; others were sense-training oriented with major emphasis directed toward eye-hand coordination and color discrimination activities; while still others were social skills oriented with major emphasis directed toward free-play activities. Yet it should be recalled that efforts were made to make the training programs in the various special classes comparable. The summer school training program, the series of in-service institutes, the curriculum guides devised by the teachers, as well as the efforts of the State Supervisor of the program were attempts in this direction. Nevertheless, the impossibility of attaining standard treatment in the various special classes must be recognized as a weakness of the study.

6. A closely related factor to comparability of training—in fact, a part of it—was differences in teacher proficiency. For generalizations to be made from the results, the teachers who instructed the experimental subjects would need to be typical of special class teachers in general. In addition, it was necessary for children of both experimental and control groups to have had comparable backgrounds; that is, no previous special

class training. Therefore, the study staff considered it highly desirable to include in the experimental group only children who were new admissions to well-established classes with experienced teachers. The number of such children made this procedure impossible, and it was necessary to draw fairly heavily on classes being established for the first time in the Fall of 1956. Of the 21 experimental subjects, 13 were enrolled with inexperienced teachers and only 8 with experienced teachers. Because this presented such a problem, a rating was made of all 30 teachers of trainable children in Tennessee. They were then rank-ordered from "most proficient" to "least proficient". (It should be pointed out that those rated lowest may still have been good teachers but were not rated as proficient as others.) The teachers were then categorized into six groups of five each in terms of "falling within the best five, falling within the second best five", and so forth. It was found that only one child of the experimental group was taught by a teacher who was ranked among the best five; eight of the children were taught by teachers who were ranked among the second best five; two were taught by teachers who were ranked among the fifth best five; and the remaining ten children were taught by teachers who were ranked among the five least efficient teachers in the State's program. It may be concluded, therefore, that in spite of the fact that steps were taken to develop common procedures and goals for the special day classes (see page 8), the assumption of "comparable" teacher proficiency was not valid.

These limitations of the study were at least partially offset by certain strengths. Some of these were as follows:

1. Perhaps the major strength of this study was the control group feature. No previous study was found in the literature where a group of "trainable" children remaining at home were identified as a control group to be compared with "trainable" children attending school.
2. The study took into account changes in parent-child relations as well as changes in the behavior of the children themselves. It was felt that special classes could be justified if they enabled the children to live more harmoniously with their families.
3. Another good feature of the study was the care taken to match the subjects on such variables as sex, etiological classification, chronological age, mental age, intelligence quotient, social age, and social quotient. This precision of matching should have reduced variance and enhanced the chances of identifying significant changes.
4. As mentioned previously, only children beginning special class training for the first time at the beginning of the experimental period were included in the experimental group. There were two reasons for this restriction. First, it was considered necessary for members of both the experimental and control groups to have had comparable backgrounds. Secondly, the study staff was of the opinion that the gains of special class training might result largely from an "initial spurt" in the development of the children resulting from their first year of placement in special classes. Every attempt was made to assure that any change which might occur would have maximum opportunity to do so.
5. As mentioned above, special care was taken in this study to develop comparable teacher proficiency. Summer school training prior to the experimental period, in-service training institutes, curriculum guides, and a state supervisor were provided in an attempt to assure

that experimental treatment would be as near optimal and comparable as practically possible.

6. A final strength of the study was in the personnel involved in developing the design, and collecting and analyzing the data. A research committee, representing a wide range of proficiencies, gave careful guidance to the research staff. Only well-trained psychological examiners with one or more years of graduate study were involved in data gathering and analyzing.

The results of the Tennessee study are supported by two major (but uncontrolled) studies undertaken by the states of Illinois and New York.

In Illinois, Goldstein[8] found appreciable changes in mental age over the first year of the project but no increase in the second year -- over and above normal rate of development which could have been predicted from I.Q. scores. In terms of improvements in self-care and socialization, teachers recorded considerable growth during the first but much less during the second year. Parents reversed this by reporting greater changes in behavior during the second year. Children with I.Q.'s of 35 and above tended to make progress in the special class while those who scored lower showed little evidence of benefiting from the program.

In New York, Johnson and Capobianco[9] reached the following conclusions at the termination of their two-year study: (1) Children with I.Q.'s below 30 will derive relatively little benefit from a day class program. (2) In general, improvements will parallel the growth curves predicted from I.Q. scores. (3) Despite the fact that small improvements will occur in habit development and social skills, the problem of the severely retarded is not resolved. Custodial care continues to be mandatory. Public school training is only a partial solution -- a total solution will involve life planning for the severely retarded individual.

Conclusions: The Tennessee Study, supported by the Illinois and New York studies provide us with two major conclusions. First, in spite of some positive trends, there is an indication that the children in the special classes did not make important amounts of progress in socialization, intellectual development, self-care, or parent-child relations over and above the children who remained at home. Second, in terms of intellectual development, the special classes as presently operated appear to produce greater changes in the high I.Q. group than the low I.Q. group. Thus it must be concluded that little evidence has accrued to support the contention that day class training for trainable children, as presently constituted, is effective. Perhaps the California study now being conducted by Cain and his associates[10] will shed new light on this problem. This investigation will involve a comparison of the development of "trainable" children in four settings: (1) community school, (2) community non-school, (3) institutional school, and (4) institutional non-school. The two new instruments being used -- the San Francisco Social Competency Scale, and the San Francisco Parent Adaptability Scale -- should be more sensitive to change than those employed in previous investigations. A major problem confronting the San Francisco study staff is the need to restrict the formal experimental period to one year.

---

[8] Herbert Goldstein, et. al. Report Number Two on Study Projects for Trainable Mentally Handicapped Children. Springfield, Ill.: State Department of Public Instruction. 1956.

[9] G. Orville Johnson and Rudolph J. Capobianco. Research Project on Severely Retarded Children. Albany, N.Y.: New York State Interdepartmental Health Resources Board. 1957.

Ultimate judgements concerning the feasibility of day classes for trainable children must be based on more variables than have been considered in studies up to this point. Such factors as the following should be taken into consideration: total psychological and economic benefits of the programs on a longitudinal basis to the children themselves, their families, and society at large; total expenses to society including human resources, physical plant, financing, etc.; the relative effectiveness of alternate programs such as residential school placement, day care centers, home training, etc. Therefore definitive statements concerning the feasibility of day classes are beyond the scope of the Tennessee study. However, it is felt that information gleaned from the Tennessee investigation warrant some recommendations:

1. The findings of the study ~~are~~ were not sufficiently conclusive to warrant discontinuance of existing day class programs for trainable children, but certain changes are suggested.

2. Serious consideration should be given to changing the criteria for selecting children to be admitted to special classes. Only children with I.Q.'s of 35 or 40 and above whose major handicapping condition is mental retardation should be included.

3. Emphasis in academic instruction should be discouraged. Unpublished data indicate that few, if any, of the trainable children are acquiring sufficient skills in the "three - R's" to warrant the attention this aspect of the curriculum is receiving. Instead, emphasis should be increased on oral language development, social adjustment, self-care, good work habits, and leisure-time activities.

It is recommended that future educational research for "trainable" children take the following directions:

1. A long-term study, similar in intensity to Terman's *Genetic Studies of Genius*, should be undertaken involving the effectiveness of the special classes for "trainable" children. This should include a follow-up of the pupils

---

[10] Leo F. Cain, et. al. *Study of the Effect of Special Day Training Classes for the Severely Mentally Retarded. Final Report of Phase I.* San Francisco, California: San Francisco State College. 1958.

into adulthood, since ultimate effectiveness of the school program, to a considerable degree, rests on how well the children adjust to home and community in later life. From such a study should come evidence for altering the special classes to make them more effective as well as provide a basis for judging their feasibility.

2. Additional study is needed of the difference in progress made in the special classes by children of different I.Q.'s with a view towards re-defining the term "trainable". By grouping together in one class children with I.Q.'s below 35-40 range and in another class those with I.Q.'s in 35-40 to 55-60 range, a comparison could be made of the relative progress of these two groups as compared with heterogeneous groups presently in existence.

3. A study is indicated on the relative effectiveness of different teachers with a view to identifying the personal qualities and training that enable some teachers to be more successful than others in training children. A study is also needed on how to increase the proficiency of teachers in the program.

4. A study is needed of mentally retarded *adults* who have not been institutionalized to identify factors which have enabled them to remain in the community.

5. Studies are suggested on the relative effectiveness of different curricula, teaching methods, physical plants, etc. The crucial question may not be whether special day class training (as it is presently constituted) is effective for trainable children (as they are presently defined). Rather it may be *what type of training* -- in terms of the abilities of the children selected for special class placement and proficiency of present teachers -- *results in the most effective school program*. This topic should be a more fruitful one for research than re-studying the broad subject undertaken in this investigation.

6. Less global studies involving the learning processes of "trainable" children are indicated. Intensive investigations of the relative effectiveness of different methods of teaching specific self-care activities provide an example.

SECTION II

Development of the "Peabody Picture Vocabulary Test"

Lloyd M. Dunn, Principal Investigator

ACKNOWLEDGEMENTS

The cooperation of many persons and institutions is essential when a test is developed and standardized on children and youth. Over 50 professional persons and 5,000 subjects participated. The investigator is indebted to the superintendents of schools, principals, supervisors of special education, school psychologists and teachers in the Nashville City and Davidson County School Systems. They graciously accepted us into their schools for purposes of examining their pupils.

Recognition is extended: Ralph T. George who served as full-time assistant to the director; John V. Hottel who carried main responsibility for deriving norms; Raymond C. Norris who consulted with the study staff on test construction and who supervised statistical procedures; Leota M. Dunn, Frances Russell, B. R. Coleman, and Mary E. Gower who administered the instrument widely in the standardization. All of the art work for the test was drawn by a commercial artist, Mrs. Sarah C. Kurek. In doing so, she made a distinct contribution by providing illustrations which expressed clearly the meaning of the stimulus words.

Last, but not least, indebtedness is acknowledged to the children who served as subjects and to their parents.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS. ii

LIST OF TABLES. v

FOREWORD. vi

Chapter

I GENERAL INFORMATION. 1

Purposes and Description. 1
Utility 1
Advantages. 2

II DIRECTIONS FOR ADMINISTRATION AND SCORING. 3

Initial Preparation by the Examiner 3
Establishing and Maintaining Rapport. 3
Introducing the Scale 4
Using the Examples. 4
Start 5
Basal and Ceiling 6
Recording Responses 6
Finding the Total Raw Score 6
Finding Derived Scores. 7
Computing the Subject's Chronological Age 7
Important Rules re: Test Administration. 7

III TEST CONSTRUCTION AND STANDARDIZATION. 10

Initial Item Selection Stage. 10
Leveling Stage. 10
Pre-testing Stage 12
Final Plate Selection Stage 14
Final Standardization 14
Reliability 18
Validity. 19

IV NORMS. 22

Age Equivalents (V.A.'s). 22
Standard Score Equivalents (V.Q.'s) 23
Percentile Equivalents (%iles). 24

LIST OF TABLES

Table		Page
1	Illustration of scoring the Peabody Picture Vocabulary Test	9
2	Summary of Item Selection and Placement.	11
3	Mean Raw Scores by Age Levels.	16
4	Standard Deviations for Raw Scores by Age Levels . . .	16
5	Reliability Coefficients for Correlating Raw Scores on Form A and Form B	17
6	Standard Errors of Measurement for Vocabulary Quotients.	17
7	Age Equivalents in Years and Months (V.A.) for Raw Scores on Form A and Form B.	25
8	Standard Score Equivalents or Vocabulary Quotients (V.Q.'s) of Raw Scores for Form A.	26
8	Standard Score Equivalents or Vocabulary Quotients (V.Q.'s) of Raw Scores for Form B.	29
9	Percentile Score Equivalents (%iles) of Raw Scores for Form A	32
9	Percentile Score Equivalents (%iles) of Raw Scores for Form B	35

APPENDIX A

Peabody Picture Vocabulary Test Individual Test Record -- Form A

APPENDIX B

Peabody Picture Vocabulary Test Individual Test Record -- Form B

APPENDIX C

Comparability of Peabody Picture Vocabulary Test Scores Under Group and Individual Administration

APPENDIX D

Peabody Picture Vocabulary Test Performance of Educable Mentally Retarded Children

APPENDIX E

Peabody Picture Vocabulary Test Performance of Trainable Mentally Retarded Children

APPENDIX F

Comparability of Peabody, Ammons, Van Alystyne and Columbia Test Scores with Cerebral Palsied Children

Foreword

An important weakness of studies of the effectiveness of day classes for "trainable" children has been lack of measuring devices receptive to small changes in the performance of the children involved. A major function of phase I of the Peabody research contract with the U. S. Office of Education was to develop a sensitive "picture vocabulary test of verbal comprehension" for use in subsequent investigations. The test was to measure verbal intelligence and hearing vocabulary. This latter function was considered especially important since educators have recognized the importance of developing oral communication among such children. The research plan called for the instrument to be especially sensitive at mental ages below six years. However it was also to extend over all levels of intellectual development so as to provide a device with sufficient top to measure those rare pupils who might be classified as mentally retarded but in fact be pseudo-feebleminded. To provide such a measure, the contract called for the instrument to be standardized on a random sample of subjects of all levels of intelligence, rather than restricting standardization solely to either the mentally retarded or the mentally normal. The two major reasons for this were as follows: (1) a better frame of reference would be provided for finding where the mentally retarded fit into the total intellectual spectrum; and (2) comparability with other tests would be increased since a number of well-accepted instruments such as the Vineland Social Maturity Scale and the Revised Standford Binet Tests of Intelligence which are frequently used in evaluating the mentally retarded were standardized on random samples from the entire population; thus by getting a derived score such as vocabulary age, test results could be compared with mental age, social age, motor proficiency age and other similar measures.

Since the test was standardized over the entire range of intelligence it should have broad utility for educational research and child study, while, at the same time, being especially applicable for investigations with "trainable" subjects. A report of the development of this instrument follows. In a sense this report serves as a manual for the test.

CHAPTER I

GENERAL INFORMATION

Purposes and Description

The Peabody Picture Vocabulary Test (PPVT) is designed to measure comprehension for the spoken word, including the ability to associate verbal symbols with pictorial representations of objects, events, and actions. It provides a well-standardized measure of (1) hearing vocabulary, and/or (2) verbal intelligence. That the test measures this latter function seems justified since vocabulary has frequently been found to be the best single predictor of school success.[1] There is growing evidence that so-called intelligence is not a unitary trait but is composed of many abilities which are present in varying amounts in different persons. Thus it is contended that the PPVT gives a rather pure measure of one of the most important aspects of intelligence - namely verbal comprehension. It should therefore predict verbal facility more precisely than numerical, mechanical, social, musical or other abilities.

The PPVT is an untimed individual test with alternate forms. Three types of derived scores are provided: (1) age equivalents (vocabulary ages or V.A.'s), (2) standard score equivalents (vocabulary quotients or V.Q.'s), and (3) percentile equivalents (%iles). Oral instructions are used as are pictorial materials, and the pointing (or "yes", "no") response. In response to a stimulus word, the subject selects from among four drawings the picture which best tells the meaning of the word. The test kit includes: (1) a spiral bound book containing a series of 150 plates preceded by three example plates which are used for both Forms A and B of the test, (2) individual record sheets listing the stimulus words and the number of the correct response for both Forms A and B, and (3) a manual of directions.

Utility

While the test was constructed primarily for use with the mentally retarded, it may be used for persons across the entire range of intelligence between the ages of 2 years 6 months through 18 years. Some success can be achieved in administering the instrument to some children below the 2 year 6 month level; for this reason the age scale norms have been extrapolated downward to the 1 year 10 months level. The 18 year norms will need to be used for adolescents and adults until normative data above the 18 year level are secured.

---

[1] There is a marked similarity between the approach used in this scale and in many tests of scholastic aptitude, reading capacity, reading readiness, and other vocabulary scales where mental ages and intelligence quotients are the derived scores.

The PPVT was constructed as a fairer test for a number of handicapped persons than many currently available. To make the test applicable to brain-injured, mentally retarded subjects with figure-ground problems of perception, the line drawings were made fairly large with fine detail excluded. Since neither an oral nor a pointing response is required of the subject, the test was made more appropriate for speech handicapped (especially the expressive aphasic and stutterer) and cerebral palsied individuals. When an oral or pointing response is not possible as with an extremely involved cerebral palsied subject, the examiner may point to the four alternates in order and elicit a "yes" or "no" from the C.P. child orally, with a head shake, or even through a coded message which may involve raising the eyelids for "yes" or closing them for "no". The PPVT has advantages in testing the non-reader, since the subject is not required to read. Too, it has advantages with certain autistic, withdrawn, and psychotic persons in that oral response on the part of the subject is not needed. This reduces the tension of the testing situation. Since the illustrations are clean, bold, line drawings, it may also be used for some partially seeing persons. In summary, the test can be given to any English speaking person between 2 years 6 months and 18 years who is able to hear words, see the drawings, and has the facility to indicate "yes" and "no" in a manner which communicates.

Advantages

The PPVT has a number of advantages. Among these are the following: (1) the test has high interest value and is therefore an easy rapport establisher; (2) extensive specialized preparation is not needed to give the test; (3) it is administered in approximately 10 to 15 minutes; (4) scoring is completely objective and quickly accomplished in one to two minutes; (5) the test covers the full range of vocabulary growth from 2 years 6 months through 18 years; (6) it is completely untimed and thus is a power rather than a speed test; (7) no verbal response is required -- a distinct advantage in testing shy and uncooperative subjects; and (8) alternate forms of the test are provided to facilitate study of vocabulary growth.

CHAPTER II

DIRECTIONS FOR ADMINISTRATION AND SCORING

Initial Preparation by the Examiner

Administration of the test requires no special preparation other than complete familiarity with this manual and practice in giving the test prior to actual use of the instrument as a standardized measure. If all the instructions contained herein are strictly observed, psychologists, teachers, speech correctionists, physicians, counselors, and social workers should be able to give the scale accurately and without difficulty. The examiner must be especially careful to acquaint himself fully with the pronunciation of each of the words to be administered. The Merriam-Webster New College Dictionary[2] has been used as the authority. In cases where two pronunciations are listed for a word, both are given to the subject. Where colloquial or other local pronunciations are accepted in a community, these alternates may be given along with the dictionary preferences.

Establishing and Maintaining Rapport

The examiner should be informal, friendly, and encouraging. Especially with pre-school, primary grade, and young handicapped subjects, the tester should give praise generously -- both when the answers are correct and when they are not. With younger children such comments as the following are usually effective: YOU'RE DOING FINE; THAT'S VERY GOOD; or MY, YOU'RE DOING WELL. With older subjects the following have been found effective: GOOD!; YOU'RE DOING AN EXCELLENT JOB. If the subject says, Did I get that one right? in most cases say: THAT IS (WAS) A GOOD ANSWER, YOU ARE DOING (DID) VERY WELL, even though an incorrect response has been made. Look and act pleased in either event. Praise can be overdone. While you should be friendly and encouraging, it is also important to be business-like especially with older subjects. Many individuals, especially older ones, will know when they are beyond their depth, and just guessing. The sensitive examiner will soom learn the optimal amount of encouragement for a particular subject to elicit the maximum cooperation and performance. In general the younger or more retarded the subject, the greater the amount of praise needed. _Remember, praise should not be confused with assistance._ There must be no direct or indirect coaching, by previous coaching on other picture material, by providing clues by glance or other behavior as to the correct illustration, or by indicating the correctness of the response.

---

[2] _Webster's New College Dictionary_ (based on Webster's New International Dictionary). Second edition. Springfield Mass.: G & C Merriam Co., 1953.

Introducing the Scale

In introducing the scale, do not mention it is a picture vocabulary or intelligence test -- unless dealing with an unusually sophisticated person. With young children (approximately 2 through 8 years) introduce the scale by calling it a "picture game", "find it game", or just a "game". Say: I WANT TO PLAY A PICTURE GAME WITH YOU. Above the eight year level, simply say: I HAVE SOME PICTURES TO SHOW YOU. With sophisticated adults, introduce the test by stating: I WANT TO FIND OUT HOW LARGE YOUR VOCABULARY IS.

Using the Examples

Preceding the 150 test plates, are three example plates. After the test has been introduced, use these three plates to establish procedure.

With average children below 8 years of age: Turn to Example A and say: SEE, THERE ARE FOUR PICTURES ON THIS PAGE. (Indicate this by pointing and saying: ONE, TWO, THREE, FOUR.) I WILL SAY A WORD, THEN I WANT YOU TO PUT YOUR FINGER ON THE PICTURE OF THE WORD I HAVE SAID. LET US TRY ONE. (Use Examples A, B, and C to make sure the subject understands the directions.) At the outset, introduce each response word by preceding it with one of the following: PUT YOUR FINGER ON ______. CAN YOU FIND ______? SHOW ME ______. POINT TO ______. FIND ______. WHERE IS ______?

With average subjects 8 years of age and over: Turn to Example A and say: YOU WILL SEE THERE ARE FOUR PICTURES ON THIS PAGE. EACH OF THEM IS NUMBERED. (Indicate this by pointing and saying: HERE ARE THE NUMBERS, ONE, TWO, THREE, FOUR.) I WILL SAY A WORD, THEN I WANT YOU TO TELL ME THE NUMBER OF THE PICTURE WHICH BEST TELLS THE MEANING OF (ILLUSTRATES) THE WORD. LET US TRY ONE. (Use Examples A, B, and C to make sure the subject understands the directions.) At the outset, introduce each response word by preceding it with one of the following: GIVE ME THE NUMBER OF THE PICTURE WHICH BEST TELLS THE MEANING OF ______. WHAT NUMBER IS ______? (Of course, subjects who prefer to point rather than call numbers should be permitted to do so.)

The examiner should use only the words given below for examples A, B, and C to illustrate the test procedure. He should be careful to select from among the examples those with which the subject is likely to be familiar. The examiner should not use as illustrations words which are beneath the subjects, for example such words as mamma and kitty which have dropped from the vocabulary of older children. Use only as many illustrations as are needed for the subject to understand the procedure; often only one example suffices.

The following are to be used in establishing the desired "pointing" or "number calling" behavior.

Pre-school children	Example A	Example B	Example C
Trial # 1	bed (1)	fish (4)	butterfly (3)
Trial # 2	spoon (4)	chair (1)	pencil (2)
Trial # 3	mamma (2)	daddy (3)	apple (4)
Trial # 4	kitty (3)	banana (2)	sock (1)
Elementary-school children			
Trial # 1	crib (1)	fins (4)	butterfly (3)
Trial # 2	silverware (4)	chair (1)	pencil (2)
Trial # 3	woman (2)	man (3)	apple (4)
Trial # 4	kitten (3)	fruit (2)	stocking (1)
Adolescents and adults			
Trial # 1	cot (1)	scales (4)	wings (3)
Trial # 2	silverware (4)	furniture (1)	eraser (2)
Trial # 3	adult (2)	carpenter (3)	stem (4)
Trial # 4	cat (3)	peeling (2)	stocking (1)

After the pointing or number calling behavior has been established, then say: THAT'S FINE. NOW I AM GOING TO SHOW YOU SOME OTHER PICTURES. EACH TIME I SAY A WORD, YOU POINT TO (TELL ME THE NUMBER OF) THE PICTURE WHICH BEST TELLS ME THE MEANING OF THE WORD. WHEN WE GET ALONG FURTHER IN THE BOOK, YOU MAY NOT BE SURE YOU KNOW THE MEANING OF SOME OF THE WORDS, BUT I WANT YOU TO LOOK CAREFULLY AT ALL OF THE PICTURES ANYWAY, AND MAKE THE BEST CHOICE YOU CAN. [The stimulus words for Form A are listed on the sheet labeled, Individual Test Record - Form A, and those for Form B on the other sheet labeled, Individual Test Record - Form B (see appendix A and B)]. Directions for Forms A and B are identical.

Start

Start subjects of estimated average learning ability with the plate numbers listed in the table below. (Subjects suspected of being slow or rapid learners may be started below or above the recommended plate in keeping with best estimates of their mental ages. Thus a twelve year old mentally retarded subject with an estimated mental age of four years would begin with plate # 15.)

Starting Points

Ages below 3-3 years	begin with plate # 1
Ages 3-3 to 4-3 years	begin with plate # 15
" 4-3 to 5-6 "	begin with plate # 25
" 5-6 to 7-6 "	begin with plate # 40
" 7-6 to 9-6 "	begin with plate # 50
" 9-6 to 11-6 "	begin with plate # 60
" 11-6 to 13-6 "	begin with plate # 70
" 13-6 to 15-6 "	begin with plate # 80
" 15-6 to 17-6 "	begin with plate # 90
" above 17-6 "	begin with plate #100

Basal and Ceiling

Work upward consecutively from the plate where the test was started until the subject makes his first error. In the event he has not made 8 consecutive correct responses prior to this first error, drop back immediately to the starting point and work backward consecutively until a total of 8 consecutive correct responses have been made by the subject. Responses above the starting point -- as well as below -- are to be counted for purposes of establishing this basal of 8 consecutive correct answers.

Continue (or recontinue) the test upward from the point of the first error until the subject makes 6 errors in any 8 consecutive presentations; count the last item presented as his ceiling. Since his ceiling has been established it marks the point where the test is to be discontinued. Thus the test is only administered over the critical range for a particular subject. (If the examiner deems it desirable he may return to easier items below the basal, after the ceiling has been established to leave the child with a success experience; however, these responses are not to enter into the calculation of the total raw score.) With some older and brighter subjects it may not be possible to establish a ceiling. In such cases the norms do not apply, and test scores should not be recorded.

Recording Responses

For each item administered, write the number of the picture which the subject indicated in the appropriate space on the individual test record form. There are at least two reasons for not just indicating correct and incorrect responses with check marks; (1) placing the number of the picture indicated will enable the examiner to re-check his scoring; and (2) the subject may see the examiner recording plus and minus, and the negative scores may result in a reduction of motivation and rapport.

For ease of scoring, indicate incorrect (wrong) responses by drawing a line through the number of the item.

Finding the Total Raw Score

The total raw score is the number of correct responses. Unanswered items below the basal point are assumed correct; unanswered items above the last item administered are assumed incorrect. To get the total raw score, subtract the number wrong from the number of the last item presented. This can conveniently be done in the spaces provided on the top portion of the individual test record forms.

Finding Derived Scores

By use of the tables found later in this report, the raw score can be converted to three derived scores: (1) an age equivalent (vocabulary age or V.A.), (2) a standard score equivalent (vocabulary quotient or V.Q.), and (3) a percentile equivalent (%ile).

Computing the Subject's Chronological Age

To compute the chronological age of the subject, use the appropriate space on the individual test record form and proceed as follows:

	Year	Month	Day
Date (date of testing)	8 195~~9~~	12 ~~1~~	50 ~~20~~
Born (date of birth)	1946	7	28
Age (chronological age)	12	6 ~~5~~	22

The subject is closer in age to 12-6 than to 12-5. Thus an extra month is added to his age. The days are not dropped unless they are 15 or less.

Important Rules re: Test Administration

1. The subject may take any reasonable amount of time per item to make his selection. However, after approximately one minute, he should be encouraged to guess. Say: GUESS. POINT TO ONE OF THEM.
2. Always secure a response. Accept guesses. Do not record "no response", or "don't know".
3. As the test proceeds, introductory phrases, i.e., Find _____. What number is _____? may be dropped with many subjects. Simply pronounce the word, and the subject will either point to his choice or announce to you its number on the plate.
4. With only a few exceptions, every word in the examples and throughout the test is in the singular, never the plural. Be sure not to convert a singular stimulus word to the plural, since this may provide a cue to the subject on certain plates where only one plural condition exists, i.e., keys.
5. Never precede the stimulus word by an article, i.e. a, an, the. This rule was established for two reasons: (1) to give consistency since articles could only be used before nouns yet a number of the stimulus words are gerunds, etc. and (2) to reduce cues since older and more able subjects could possibly reject certain of the decoys by knowing the correct selection must depict primarily an object rather than an action.

6. For very young children, i.e. normal boys and girls below 3 years of age and severely retarded children, the examiner may have to establish the subject's pointing response. To do this say: PUT YOUR FINGER ON ______, at the same time placing his finger on the correct picture. After a few such trials the examiner may take the lead by pointing and then encouraging the subject to do likewise. The length of time to establish the desired pointing behavior will vary from child to child. However, the subject should be able to take over after the three example plates have been used. In very difficult cases the examiner may wish to repeat the examples. If the pointing behavior has not been established at this point, it is obvious that the test should be discontinued.

7. Some of the subjects may perseverate and begin pointing to one corner of the plate time after time. The examiner should be on the watch for this, and get the subject to look at all pictures before pointing out his response. This can be done by the examiner pointing to picture #1 and saying: LOOK AT THIS ONE; then pointing to #2 and saying: AND THIS ONE; then to #3 and saying: AND THIS ONE; then to #4 and saying: AND THIS ONE.

8. Especially with younger children, it is necessary to repeat frequently: BE SURE TO LOOK CAREFULLY AT ALL FOUR PICTURES, or simply: LOOK AT ALL OF THE PICTURES.

9. It is not permissible to define or spell stimulus words, or use them in sentences.

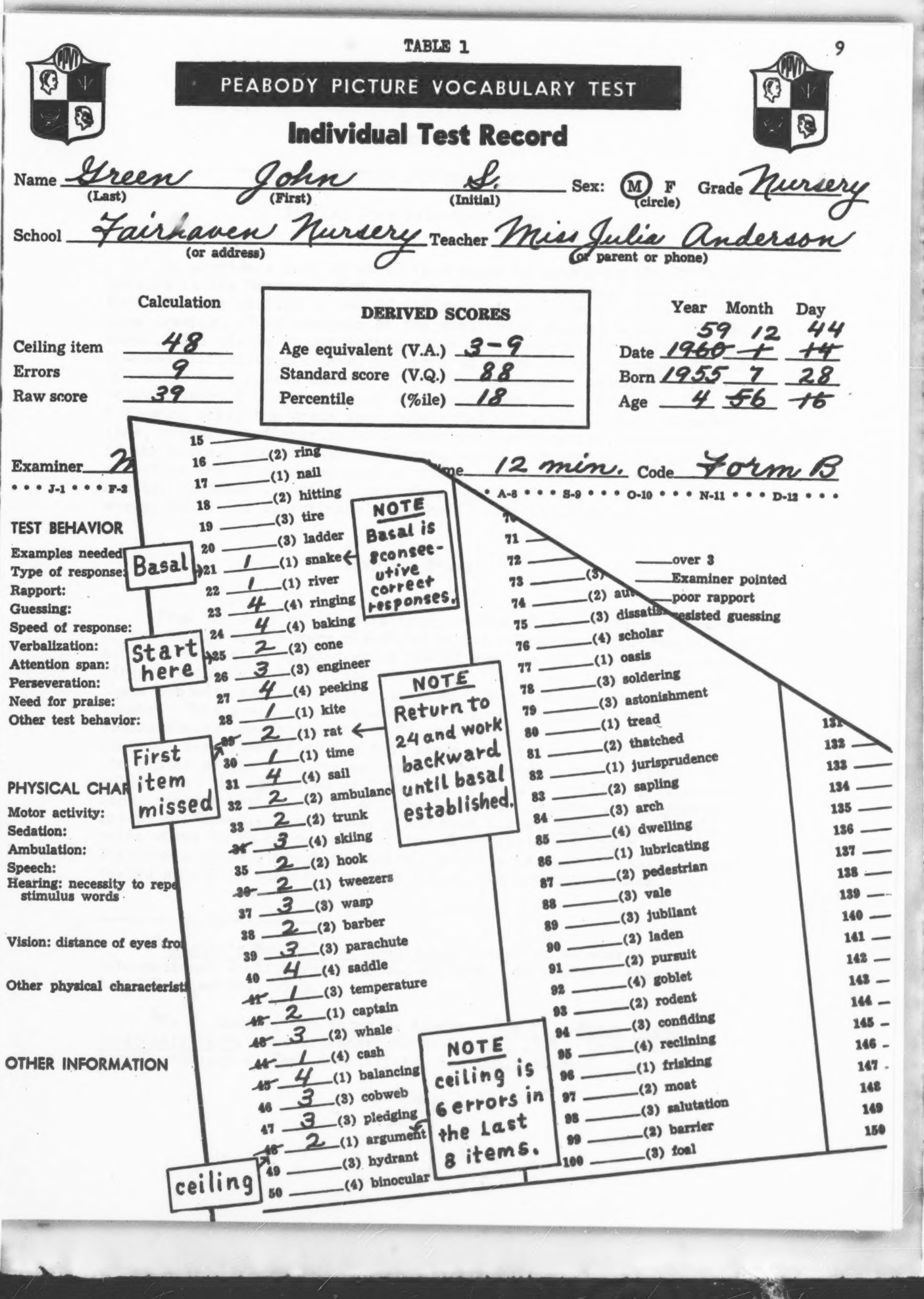
10. It is not permissible to show the subject the printed stimulus word.

11. Stimulus words may be pronounced aloud more than once by the examiner.

Note

See Table I for an example of how to complete the Individual Test Record.

TABLE 1



PEABODY PICTURE VOCABULARY TEST

Individual Test Record

Name Green John S. Sex: [✓ M] F Grade Nursery
(Last) (First) (Initial) (circle)

School Fairhaven Nursery Teacher Miss Julia Anderson
(or address) (or parent or phone)

Calculation

Ceiling item 48
Errors 9
Raw score 39

DERIVED SCORES

Age equivalent (V.A.) 3-9
Standard score (V.Q.) 88
Percentile (%ile) 18

	Year	Month	Day
	59	12	44
Date	~~1960~~	~~1~~	~~14~~
Born	1955	7	28
Age	4	~~5~~6	~~16~~

Examiner

Time 12 min. Code Form B

• • • J-1 • • • F-2 • • • A-8 • • • S-9 • • • O-10 • • • N-11 • • • D-12 • • •

TEST BEHAVIOR

Examples needed
Type of response:
Rapport:
Guessing:
Speed of response:
Verbalization:
Attention span:
Perseveration:
Need for praise:
Other test behavior:

over 3
Examiner pointed
poor rapport
resisted guessing

PHYSICAL CHAR
Motor activity:
Sedation:
Ambulation:
Speech:
Hearing: necessity to repe stimulus words
Vision: distance of eyes fro
Other physical characterist

OTHER INFORMATION

15
16 (2) ring
17 (1) nail
18 (2) hitting
19 (3) tire
20 (3) ladder
21 1 (1) snake
22 1 (1) river
23 4 (4) ringing
24 4 (4) baking
25 2 (2) cone
26 3 (3) engineer
27 4 (4) peeking
28 1 (1) kite
~~29~~ 2 (1) rat
30 1 (1) time
31 4 (4) sail
32 2 (2) ambulanc
33 2 (2) trunk
~~34~~ 3 (4) skiing
35 2 (2) hook
~~36~~ 2 (1) tweezers
37 3 (3) wasp
38 2 (2) barber
39 3 (3) parachute
40 4 (4) saddle
~~41~~ 1 (3) temperature
~~42~~ 2 (1) captain
~~43~~ 3 (2) whale
~~44~~ 1 (4) cash
~~45~~ 4 (1) balancing
46 3 (3) cobweb
47 3 (3) pledging
~~48~~ 2 (1) argument
49 (3) hydrant
50 (4) binocular

Basal

NOTE
Basal is 8 consecutive correct responses.

Start here

NOTE
Return to 24 and work backward until basal established.

First item missed

NOTE
ceiling is 6 errors in the last 8 items.

ceiling

70
71
72
73 (3)
74 (2) au
75 (3) dissatis
76 (4) scholar
77 (1) oasis
78 (3) soldering
79 (3) astonishment
80 (1) tread
81 (2) thatched
82 (1) jurisprudence
83 (2) sapling
84 (3) arch
85 (4) dwelling
86 (1) lubricating
87 (2) pedestrian
88 (3) vale
89 (3) jubilant
90 (2) laden
91 (2) pursuit
92 (4) goblet
93 (2) rodent
94 (3) confiding
95 (4) reclining
96 (1) frisking
97 (2) moat
98 (3) salutation
99 (2) barrier
100 (3) foal

131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150

CHAPTER III

TEST CONSTRUCTION AND STANDARDIZATION

Initial Item Selection Stage

To provide a pool of words from which to construct the PPVT, all entries in the Merriam Webster's New College Dictionary[3] were examined. Each word was studied to see if its meaning could be illustrated by a line drawing. This analysis of the dictionary yielded a pool of 3,885 words. The staff, by _subjective_ consensus, tentatively arrived at a guessed difficulty age level for each of these words. Between 80 and 350 words were placed at each of the 24 age levels, extending from below 2 years 6 months through adulthood (see Column #2 Table 2). Following this, the words were catalogued and a search made of magazines, picture dictionaries, and other publications for illustrations of the words. When it was considered especially appropriate to include a word in the illustrated pool of items, yet no picture could be found, artists sketched illustrations. From the dictionary pool of 3,885 words, 2,055 (or approximately one-half) were illustrated.

Leveling Stage

From the pool of illustrations, groups of six pictures at each of the 24 age levels were placed on series of cards. Approximately 16 subjects were then located whose mental ages on the Ammon's Full-Range Picture Vocabulary Test[4] fell at each of the 24 age levels. (All of these subjects resided outside of Nashville proper but in Davidson County. This precaution was necessary since the plan called for standardizing the instrument on Nashville subjects.) In all 360 subjects were identified for this stage of the test construction. The plates of six pictures each were then administered to groups of subjects at the estimated critical age levels. By shifting illustrations, reconstructing plates and re-administering them to the appropriate groups of "leveling" subjects, items were placed _objectively_ at the point where 40 to 60 per cent of the appropriate age group passed. This was a gross (though objective) leveling in that the illustrations varied widely in quality, clarity, size, and appropriateness. In a sense, at this stage in the test construction there were six forms to the test, since each illustration represented a word, and the plate was re-administered six times to the appropriate group of subjects to check on the difficulty level of each word illustrated on the plate. This procedure resulted in placing between 57 and 196 words at each of the 24 age levels for a total of 2,055 words (see Column #3, Table 2).

---

[3] _Webster's New College Dictionary_ (based on Webster's New International Dictionary). Second edition. Springfield Mass.: G & C Merriam Co., 1953.

[4] R. B. Ammons, and Helen S. Ammons. _The Full-Range Picture Vocabulary Test_. New Orleans: R. B. Ammons, 1948. (now available from Psychological Test Specialists, Box 1441, Missoula, Montana)

TABLE 2

SUMMARY OF ITEM SELECTION AND PLACEMENT

	Number of words			Number of plates	
Age levels	Initial item selection stage	Leveling stage	Pre-testing stage	Pre-testing stage	Final plate selection stage
1	2	3	4	5	6
Adult #1	155	57	32	8	7
Adult #2	170	60	32	8	8
Adult #3	175	68	32	8	9
18-0	100	75	32	8	10
17-0	110	70	32	8	4
16-0	125	75	32	8	4
15-0	240	80	32	8	4
14-0	200	77	32	8	5
13-0	175	77	32	8	5
12-0	350	85	32	8	5
11-0	200	78	32	8	5
10-0	240	107	32	8	6
9-0	185	70	32	8	6
8-0	225	68	32	8	6
7-0	160	121	32	8	5
6-0	150	100	32	8	4
5-6	100	94	32	8	4
5-0	80	64	32	8	4
4-6	85	71	32	8	4
4-0	120	97	32	8	6
3-6	115	107	32	8	6
3-0	95	77	32	8	7
2-6	110	81	32	8	8
Below 2-6	240	196	64	16	20
24	3,885	2,055	800	200	150

2) Number of words gleaned from the dictionary and subjectively leveled by the study staff.
3) Number of words objectively leveled (40-60 % passing at a given age level).
4) Number of words illustrated by the artist (and used in pre-test plates).
5) Number of pre-test plates.
6) Number of plates in the final series.

Pre-testing Stage

At this point, plates similar in format to those used in the final battery were constructed with groups of four items selected to make up each plate.

Criteria for the selection of the four words to be used in making up a plate were:

1. All four words were found at the leveling stage to be of the same difficulty level.
2. All four words demonstrated good linear growth curves in terms of per cent passing at successive age levels. (Words with flat curves, irregular curves, etc., were excluded.)
3. Words were fixed at the age level where 40 to 60 per cent of the subjects of a given age level passed the item.
4. Words were used where no sex differences were found to exist.
5. Words were omitted which were biased culturally, regionally, and racially. Avoided also were dated words, plurals, double words, scientific terms, etc.
6. Only singular and collective words were used.
7. Primarily nouns, some gerunds, and a few adjectives and adverbs were used.

Criteria for selection of decoys changed from the very early to the later plates. In the first few plates, illustrations were kept very dissimilar to reduce the difficulty of the plate. For example, only one human figure was used in any of the early plates. Plate #1 of the final battery will illustrate the procedure: a human form, a means of locomotion, a piece of furniture, and a toilet utensil were used. Later in the test the plates were made more difficult by making the decoy illustrations very similar to the response item, e.g. in Plate #75, the correct response and decoy drawings are all human forms. Similarly in Plate #80 all four illustrations are of manufactured objects, and in Plate #150 all are plant and animal life. Categories for plates included: man-made objects, animals, birds, human action (gerunds), nature scenes, plants, flowers, inanimate objects, adverbs, articles in a home, adjectives, musical instruments, occupations, scientific materials, parts of a house, and wearing apparel.

As will be seen from the four examples below, the four words selected for any one plate were: (1) of the same difficulty, (2) of the same part of speech, (3) somewhat equal in length, (4) began with different speech sounds, and (5) except for the plate #1, were usually similar in category.

Plate number	Picture Number 1	2	3	4
1	brush	table (b)*	doll (c)*	car (a)*
75	startle (c)	concentration	dissatisfaction (b)	destruction (a)
80	tread (b)	coil (a)	projector	valve (c)
150	falcon	nidificate (c)	pensile (a)	culver (b)

* Letters designate words in Form A, B, and C of the pre-test battery.

When the plates were made up, they were then turned over to the commercial artist for illustration. (It will be recalled that clippings from magazines etc., were used as illustrations at the leveling stage.) The illustrations were redrawn according to the following criteria: (1) equal size, (2) equal intensity, (3) equal appeal, and, (5) appropriate to the age level of subjects most likely to view the plate in the test situation, i.e. earlier plates were in general of young children, middle numbered plates of elementary children, and later plates were of adolescents and adults. The drawings were made large and clear with simple line drawings which were relatively free of figure-ground and other distracting elements.

A total of 8 plates were designed at each of 23 age levels, while 16 plates were developed at the level below the 2 years 6 month. (This was done to make the test especially sensitive for testing very young and mentally retarded children.) Therefore, 32 words were illustrated at all but the lowest level where illustrations were made for 64. Thus the artist illustrated a total of 800 words (see Column #4 and 5, Table 2).

As will also be seen from the four examples at the top of the page, at this pre-test stage three forms of the test existed--Forms A, B, and C--and if time had permitted a fourth, Form D, could have been used since all four words were of approximately the same difficulty level etc. This additional Form C was to allow for substitutions without re-drawing the plates for the final form of the test. A word designated as Form A on the pre-test battery could remain a Form A word on the final battery, become a Form B word, or be dropped from the test entirely.

When the 200 plates (800 illustrations) were drawn by the artist, they were turned over to the printer to be reduced to 8" x 10½" prints. Thus the illustrations used in the pre-testing were exactly the same as those found in the final battery.

All three forms of the pre-test battery were administered to 750 subjects in Davidson County, Tennessee. Below the 8 year level it was administered individually; from the 8 year level and on it was administered to groups of subjects through projecting slides of the plates on a screen. (See Appendix C)

Final Plate Selection Stage

A careful item analysis was made of the 800 words administered through the 200 pre-test plates. Curves were drawn for all three stimulus words (Forms A, B, and C) on each plate to indicate per cent passing at successive age levels. Also calculated was the per cent of subjects selecting each of the decoys. These data enabled the staff to make up the final test battery of 150 plates. Words were only included when the curves indicated a steady and increasing per cent passing from one age level to the next. Words used as Form A, B, and C in pre-testing were re-shuffled to make up the final Form A and B. The two words retained were those most equal in difficulty and with the best curves of increase in per cent passing with increase in age. These two words were assigned to Forms A and B in such a way as to equate difficulty of the forms for each year level. Some plates were completely rejected at this stage. In others, when a decoy was not working, that decoy was redrawn or a substitution made, and the plate was pre-tested again. Because of the improved artist-drawn illustrations in the pre-test battery as compared with magazine clippings in the leveling battery, the plates shifted considerably, and generally downward. Thus an even pool of usable plates was not available at each level. Because of this, it was necessary to make up the final battery of 150 plates with less than equal and ideal numbers at each age level (see Column #6, Table 2).

Final Standardization

The PPVT was standardized on 4,012 cases as follows:

Age level	No. of cases	Age level	No. of cases
18	227	8	100
17	305	7	100
16	354	6	183
15	307	5	133
14	269	4-6	122
13	287	4-0	110
12	250	3-6	119
11	384	3-0	92
10	319	2-6	92
9	259		
		Total	4,012

All subjects were administered both Forms A and B of the final battery during the period April-June, 1958. The alternate forms were counterbalanced for order of presentation. Not less than three and not more than seven days elapsed between administrations. Only white children and youth residing in Nashville, Tennessee were included in the final standardization group. However, certain precautions were taken to provide norms which should be useful throughout the United States. Different procedures were used for subject selection and/or test administration at the (1) pre-school, (2) early elementary, and (3) upper-elementary and high school levels.

1. At the pre-school level (2 years 6 months through the lower half of the 6 year level) the following procedure was used. The Kuhlmann-Finch Intelligence Tests[5] are administered to all pupils in the Nashville City Schools periodically beginning with the Grade three level. The school psychologist of Nashville had prepared I.Q. score distributions by schools on this test. A sample of four elementary schools were selected which provided a composite normal probability curve of I.Q.'s for the pupils enrolled in these schools. The geographical areas served by these four schools were then used to draw the sample of pre-school subjects. Principals and teachers of the four schools involved conducted a census of all pre-school children in their school zones. They then requested permission of the parents to allow the four psychological examiners involved in this aspect of the standardization to visit the homes where pre-school children lived for the purpose of administering the PPVT. The homes were visited on two occasions so that both Forms A and B of the test could be administered individually to all subjects.

2. At the lower elementary school level (upper half of the six year level as well as the 7 and 8 year levels) the following procedure was used. The Metropolitan Readiness Tests[6] are administered to Grade one children and the Kuhlmann-Finch Intelligence Tests to Grade three children in the Nashville City Schools. Random samples of Grade one, two and three boys and girls were drawn from the classes of the Nashville School System until a normal distribution of scores was approximated. These subjects --50 boys and 50 girls at each of the 6½, 7, and 8 year levels -- were then given both Form A and B of the PPVT individually. Four schools were used.

3. At the upper elementary and high school levels (ages 9 through 18 years) the following procedure was used. Here the PPVT was administered as a group test by the use of photographic slides of the series of plates, a semi-automatic projector, a screen, and an especially-designed answer sheet. At least three persons proctored each administration: One operated the projector, another read off the words, and the third supervised the subjects. Prior to adopting this procedure, an experiment was conducted to determine if scores obtained on the PPVT would differ significantly under individual and group administrations. (The report of the study appears in Appendix C.) No significant or appreciable differences were found in PPVT scores obtained under group and individual administrations. Because of this evidence the test was administered by the group technique to children and youth from the 9 year level and on. Thus the size of the norm sample was increased appreciably accompanied by a considerable saving of time and expense. Four elementary schools, two junior high schools, and two senior high schools were used.

---

[5] Frank H. Finch. Kuhlmann-Finch Intelligence Tests. Minneapolis: American Guidance Service, 1953.

[6] Gertrude H. Hildreth and Nellie L. Griffiths. Metropolitan Readiness Tests. Yonkers-on-Hudson, N.Y.: World Book, 1950.

TABLE 3

MEAN RAW SCORES BY AGE LEVELS

		Mean Scores					
		Form A			Form B		
		Boys	Girls	Total	Boys	Girls	Total
	18-0	110.66	107.31	108.99	112.18	108.96	110.57
	17-0	106.79	103.58	105.11	108.86	104.53	106.60
	16-0	102.64	101.85	102.25	102.53	102.89	102.70
	15-0	98.63	95.07	96.88	99.72	98.45	99.10
A	14-0	94.33	92.37	93.36	95.12	94.93	95.03
G	13-0	91.32	90.06	90.69	92.58	91.35	91.96
E	12-0	86.41	83.07	84.62	87.60	85.04	86.23
	11-0	83.92	80.12	82.23	84.88	81.64	83.22
L	10-0	76.76	74.30	75.49	77.29	75.02	76.11
E	9-0	74.70	68.52	71.29	74.81	68.85	71.52
V	8-0	66.38	65.24	65.81	66.10	65.74	65.92
E	7-0	62.12	59.52	60.82	61.78	59.72	60.75
L	6-0	55.09	55.68	55.37	55.61	54.74	55.20
	5-0	51.17	49.00	50.23	50.53	48.43	49.62
	4-6	45.86	45.26	45.58	45.42	44.74	45.05
	4-0	43.44	41.21	42.08	43.49	41.07	42.02
	3-6	37.07	34.43	35.67	36.30	34.14	35.16
	3-0	30.75	26.79	29.28	30.76	27.85	29.68
	2-6	18.53	21.94	20.39	18.62	21.58	20.23

TABLE 4

STANDARD DEVIATIONS FOR RAW SCORES BY AGE LEVELS

		Standard Deviations					
		Form A			Form B		
		Boys	Girls	Total	Boys	Girls	Total
	18-0	15.12	14.69	14.97	15.31	14.09	14.77
	17-0	13.73	14.36	14.13	13.97	13.70	13.98
	16-0	14.21	14.36	14.27	13.48	12.71	13.09
	15-0	12.89	13.82	13.45	13.54	13.85	13.65
A	14-0	12.51	11.40	11.99	13.67	13.02	13.33
G	13-0	11.78	11.57	11.67	12.84	12.57	12.70
E	12-0	10.70	9.51	10.20	10.46	10.97	10.79
	11-0	9.62	9.94	9.95	10.99	10.60	10.90
L	10-0	9.77	9.39	9.64	9.30	9.18	9.29
E	9-0	8.84	8.23	9.03	9.25	8.44	9.29
V	8-0	8.77	8.49	8.60	9.39	8.02	8.69
E	7-0	7.43	7.95	7.77	7.41	7.66	7.61
L	6-0	7.38	7.70	7.52	7.64	6.24	7.00
	5-0	7.79	8.54	8.17	8.43	7.44	8.05
	4-6	9.34	7.85	8.64	8.23	7.20	7.74
	4-0	10.19	10.41	10.34	8.27	9.06	8.80
	3-6	9.80	9.56	9.72	10.21	8.67	9.45
	3-0	9.79	9.05	9.66	9.57	7.05	8.80
	2-6	6.70	9.47	8.45	6.96	9.54	8.54

TABLE 5

RELIABILITY COEFFICIENTS FROM CORRELATING RAW SCORES ON FORM A AND FORM B

		Correlations Forms A and B		
		Boys	Girls	Total
AGE LEVEL	18-0	.87	.82	.84
	17-0	.83	.85	.84
	16-0	.78	.82	.80
	15-0	.80	.85	.83
	14-0	.84	.77	.81
	13-0	.81	.59	.70
	12-0	.77	.79	.78
	11-0	.78	.82	.81
	10-0	.77	.76	.77
	9-0	.73	.68	.74
	8-0	.84	.73	.79
	7-0	.66	.79	.74
	6-0	.76	.57	.67
	5-0	.74	.71	.73
	4-6	.74	.68	.72
	4-0	.79	.76	.77
	3-6	.82	.81	.81
	3-0	.75	.73	.75
	2-6	.65	.80	.76

TABLE 6

STANDARD ERRORS OF MEASUREMENT FOR VOCABULARY QUOTIENTS

		Standard Errors		
		Boys	Girls	Total
AGE LEVEL	18-0	5.42	6.36	6.00
	17-0	6.18	5.80	6.00
	16-0	7.04	6.36	6.70
	15-0	6.70	5.80	6.18
	14-0	6.00	7.20	6.54
	13-0	6.54	9.60	8.22
	12-0	7.20	6.87	7.04
	11-0	7.04	6.36	6.54
	10-0	7.20	7.35	7.20
	9-0	7.80	8.49	7.65
	8-0	6.00	7.80	6.87
	7-0	8.74	6.87	7.65
	6-0	7.35	9.84	8.61
	5-0	7.65	8.08	7.80
	4-6	7.65	8.49	7.94
	4-0	6.87	7.35	7.20
	3-6	6.36	6.54	6.54
	3-0	7.50	7.80	7.50
	2-6	8.88	6.70	7.35

Reliability

An important characteristic of a test is its reliability -- the degree to which a subject's score remains stable or consistent from one form of the test to another, and from one time to another. Reliability coefficients were obtained from scores of the standardization subjects. Pearson product-moment correlations were calculated between raw scores on Form A and Form B for each sex separately, for each age level (see Table 5). In gathering these data, the forms were administered in counterbalanced order with three to eight days separating administrations. This interval was selected for two reasons. To allow for day to day variability in test performance, the alternate forms were not administered simultaneously. To hold the vocabulary of the subjects relatively constant, only a few days were allowed to elapse between test and retest.

Correlations ranged from a low of 0.67 at the 6 year level to a high of 0.84 at the 17 and 18 year levels when the scores of both males and females were combined.

Standard errors of measurement for standard scores (V.Q.'s) were calculated from the parallel forms reliability coefficient. The formula used was: $S.E._m = \pm 15 \sqrt{1 - r_{12}}$. The resulting standard errors of measurement are listed in Table 6. When the scores of males and females were combined by age levels, the standard errors ranged from a high of 8.61 points at the 6 year level to a low of 6.00 points at the 17 and 18 year levels.

The evidence based on the standardization data indicates that the PPVT is a reliable instrument within single age levels, and to a considerable degree for each sex. Supportive reliability information is available.

Norris, Hottel, and Brooks examined the scores of 60 fifth grade pupils of average intelligence, concluding from Form A vs Form B comparison that differences in raw scores were "so small as to leave little doubt as to the equivalence of the two forms . . . . Furthermore, any practice effect which might have stemmed from having had the alternate form two days earlier averaged only about one point" (see appendix C).

Dunn and Brooks administered both forms of the PPVT to 130 "educable" mentally retarded subjects, ages 7 to 16 years. The reliability coefficient for age equivalent scores was 0.76 (see appendix D).

Dunn and Hottel administered the alternate forms of the PPVT to 267 "trainable" mentally retarded subjects, ages 7 to 16 years. The reliability coefficient for age equivalent scores was 0.87 (see appendix E).

Dunn and Harley found an alternate form reliability coefficient of 0.97 for 20 cerebral palsied children of ages 7-1 through 16-2 years.

This evidence suggests the PPVT is a reasonably stable instrument for average, mentally retarded, and cerebral palsied subjects. No evidence has been accumulated on test-retest or long-term reliability of PPVT scores.

Validity

Another important characteristic of a test is its validity -- the extent to which it measures what it purports to measure. This includes validity information for both individual items and for the total test. There are two main types of validity evidence -- rational and statistical.

Two common types of rational validity are content and construct validity. The PPVT has high content validity in that a complete search was made of the Merriam Webster New College Dictionary for words which could be illustrated. The restriction was omitting words which could not be illustrated. Since a good cross section was obtained of words in common use today in the United States, and since care was taken to keep the final selection of response and decoy items unbiased, the final product is assumed to meet adequate standards of test construction for content validity.

As long as the PPVT is assumed to measure hearing vocabulary, its rational validity rests on its content validity. However, when an inference is made that the test measures verbal intelligence or scholastic aptitude, rational validity must be based on construct validity. Again the PPVT should have reasonably good validity in that the items used have been found to predict well the concept or construct which the test if designed to measure. For the Revised Stanford-Binet Tests of Intelligence[7], Terman and Merrill have stated: "We have found the vocabulary test to be the most valuable single test in the scale". For the Wechsler Intelligence Scale for Children[8], Wechsler found the vocabulary sub-test scores to correlate more highly with Full scale I.Q. scores than any other sub-test. Numerous other studies investigating the measurement of intelligence have shown that vocabulary is the best single item for prediciting school success[9]. The PPVT purports to measure verbal intelligence rather than social, practical, or mechanical intelligence. It is believed the test will predict language arts achievement better than achievement in science and mathematics, though this remains to be demonstrated.

---

[7] Lewis M. Terman and Maud A. Merrill. Measuring Intelligence. Boston: Houghton Mifflin, 1937. p. 302.

[8] David Wechsler. Wechsler Intelligence Scale for Children Manual. New York: Psychological Corporation, 1949. p. 10.

[9] Edgar Dale and Donald Reichert. Bibliography of Vocabulary Studies. Columbus, Ohio: Ohio State University, Bureau of Educational Research, 1957.

Common types of statistical validity include item validity, congruent validity, concurrent validity, and predictive validity. As yet, the major case for the PPVT rests on item validity. Validity of each item was judged by increase in per cent passing from one age level to the next, based on plotted curves for each item. Fairly rigorous standards were maintained through extensive leveling and pre-testing of items. In the final pre-testing alone a total of 600 stimulus words (200 plates x 3 forms) were reduced to 300 stimulus words (150 plates x 2 forms) for the final test battery. Further evidence of validity is found in Table 3. Here mean raw scores on the PPVT for Forms A and B are shown for succeeding age levels. As will be seen, mean raw scores increased with age, in a linear fashion, from 1-6 to 18-0 year levels. This should not be construed as evidence that the total vocabulary of children and youth increases in a similar fashion. In fact, the test was constructed with a large number of items at the early pre-school level to correct for anticipated curvilinearity. Increase in raw scores with successive ages is indirect but not conclusive evidence of validity since physical strength, ossification, height, and weight also increase with age yet are not highly related to verbal intelligence.

A test has congruent validity when it compares well with other well-established measures of the same function. Evidence of congruent validity was obtained only on mentally retarded subjects wherein their scores on the PPVT were correlated with scores on the Revised Stanford-Binet Tests of Intelligence. On the 130 "educable" children ages 7 to 16 years, the validity coefficient for vocabulary age vs mental age was 0.74 (see appendix D). On the 267 "trainable" children, ages 7 to 16 years, a coefficient calculated in the same fashion was 0.71 (see appendix E). It was not anticipated that the PPVT would correlate perfectly with other tests of intelligence. The author believes the PPVT and Binet overlap in the function(s) they measure but do not correspond perfectly. In fact, research may show the PPVT to correlate more highly with language arts achievement than the Binet, and the Binet to predict more completely over-all school achievement including mathematics and science. Correlatingthe PPVT with other intelligence tests probably has limited utility, and involves circularity in reasoning because it does not indicate whether the one test is better or worse than the other.

Concurrent validity was found by correlating PPVT scores with teacher ratings of language arts ability obtained at the same time the test was administered. With 309 "trainable" subjects the correlations between teacher ratings of reading achievement and vocabulary age scores was 0.52. Data on 20 cerebral palsied children yielded a correlation of 0.90 between Form A scores and teacher ratings of arithmetic achievement and, 0.87 between Form A scores and teacher ratings of reading achievement. (Results for Form B were very similar.)

No predictive validity data are yet available on the PPVT. In fact all of the statistical validity on the test are limited and preliminary. As with other new instruments, the validity evidence accumulates after it reaches the field and comes under the scrutiny of researchers.

Successive editions of this manual will contain this accumulated evidence. Its weakest point probably is at the superior adult level, since vocabulary has been found lacking in discriminating between such individuals. At this level, eduction of correlates and analogies have been found to have higher predictive validity for scholastic success. Another area of weakness appears to be at the 5-0 to 6-0 year levels. As seen in Table 2, insufficient words remained at these levels to construct the desired number of plates for the final battery. As seen in Table 5, the reliability coefficient was lowest at the 6-0 year level.

CHAPTER IV

NORMS

Three types of norms have been derived for the PPVT: age equivalents (V.A.), standard score equivalents (V.Q.), and percentile equivalents (%ile).

Age Equivalents (V.A's)

Age norms provide an index of the level of development for a given subject. For example, 75 is the mean raw score on Form A for children who have chronological ages of 10-0 (see Table 7). Therefore, regardless of the subject's chronological age, if he obtained a raw score of 75 on Form A of the PPVT, he is said to possess a vocabulary level of ten years since his ability to score on the PPVT is like that of the average ten year old. Approximate grade equivalents are derived from age equivalents by the rule of five. Thus a child with a V.A. = 10-0 would have a grade equivalent of 5.0 (subtract 5 from the V.A.). If one accepts the assumption that the PPVT measures verbal intelligence then age equivalents are comparable to mental age scores. Then a vocabulary age of 10-0 would indicate the subject has accumulated the capacity to achieve at the beginning grade 5 level, at least in the language arts.

Age norms have the advantages of giving us an easily understood index of a subject's developmental level in vocabulary growth; they are useful in comparing vocabulary age with chronological age, achievement age, and so on. In addition they provide meaningful derived scores for the extreme deviate. Standard score equivalents were only derived out to three standard deviations. Thus the lowest and highest V.Q. scores are 55 and 145 respectively. For subjects who score below 55 or above 145, age norms are the most useful derived scores available.

Age norms for converting raw scores on the PPVT to vocabulary age scores are shown in Table 7. These are derived by the Mean Test Age method. From the data in Table 3 separate graphs were plotted for Forms A and B of chronological age vs mean raw score at each age level 2-6 through 18-0 years. To accomplish this, a graph was first constructed on which all possible PPVT raw scores were distributed along the Y-axis and chronological age in one-month intervals was distributed along the X-axis. Then the mean raw score for each age level was plotted on this graph. These plots were then connected by straight ruled lines and an age-score curve resulted. i.e. simple linear interpolation. Age equivalent scores were then derived by reading chronological age to the closest month from the point where raw-score lines intersected the age-score curve.

Since the mean raw scores for subjects 2-6 was 20, age equivalents for raw scores below 20 were obtained by extending the curves and taking age readings from these extrapolations. Thus age scores below 2-6 should be considered as estimates only.

Standard Score Equivalents (V.Q.'s)

Standard score norms (V.Q.'s) provide an index of brightness for a given subject in comparison with other individuals of the same age. A V.Q. of 100 on the PPVT was arbitrarily assigned to the mean raw score for subjects at each age level. The standard deviation was set at 15 V.Q. points. This provided the following classification:

V.Q.	Per cent included	Classification
125 and above	5	Very rapid learners
110 to 125	20	rapid learners
90 to 110	50	average learners
75 to 90	20	slow learners
Below 75	5	very slow learners

Standard scores have the advantage of enabling the examiner to compare a subject's test score with the scores made by other individuals in his age group. They provide scores which can easily be handled statistically. Too, they are not based on assumptions of linearity of vocabulary growth with chronological age growth as would have been the case had V.Q.'s been derived by the formula V.Q. = V.A./C.A. Further, the problem of keeping V.A. units equal with adolescents and adults did not have to be encountered. Finally, since the standard deviation was kept constant from year to year, a subject's V.Q. would not be expected to vary unless his score on the test relative to others in his own age changed.

The standard score norms for converting raw scores to vocabulary quotients are shown in Table 8. These were derived by preparing separate raw score distributions for each age level in the standardization sample. The mean and standard deviation of raw scores was then found for each distribution. The raw score distribution at each level was then converted to a standard score scale by the following formula: V.Q. = 100 + (raw score deviation from 100 x 15 / standard deviation of the raw score distribution). Thus it can be seen that the standard score distribution for each age level is based upon the actual distribution of raw scores for that age level in the normative group.

As mentioned above, standard score equivalents were derived only out to 3 standard deviations. Thus in Table 8 the lowest V.Q. is 55 and the highest 145. (Only an estimated 0.2 per cent of the population fall above and below this point.)

The vocabulary quotients toward the extremes of the age distributions are irregular. This resulted because no "smoothing" was made in score distributions. The decision to let the data speak for themselves is reflected in slight reversals at the extremes of the derived score distributions.

Percentile Equivalents (%iles)

Percentile norms also provide an index of brightness, indicating how a given individual compares with others of his own age. Thus a person at the 50 %ile, has scored in such a fashion that, for every 100 persons at a given age level on whom the test was standardized, 50 scored above and 50 scored below him. A person at the 25 %ile has 25 individuals of his age scoring below him for every 75 scoring higher than he did.

Percentiles have the disadvantage (as compared with standard score norms) of providing unequal units. Thus the difference between the 5th and 10th percentile is far greater than between the 55th and 60th percentile. It has one major advantage over standard score norms in that no assumption is made about the normality of distribution of scores at any age level.

The percentile norms for converting raw scores to percentiles are shown in Table 9. These were derived by the method set forth in Walker. (Walker, Helen M. Elementary Statistical Methods. New York: Henry Holt, 1943, p. 64.) Working upward from the lowest score in each age level, the percentage of scores equaled or exceeded by a given score was computed by the formula:

$$\text{Percentile (for a given raw score)} = \frac{1}{N}\left(\frac{Fi}{2} + CF\right)$$

Where N = number of cases in distribution, Fi = frequency within an interval, and CF = cumulative frequency below the interval.

As in the case of standard score equivalents, the percentile distributions were not smoothed. Thus some slight reversals between age levels will be noted in the extremes of the %ile norms.

TABLE 7

AGE EQUIVALENTS IN YEARS AND MONTHS (V.A.) FOR RAW SCORES ON FORM A AND FORM B

Raw Scores	Vocabulary Age Form A	Vocabulary Age Form B	Raw Scores	Vocabulary Age Form A	Vocabulary Age Form B
111	-	18-0	57	6-3	6-4
110	-	17-10	56	6-1	6-2
109	18-0	17-7	55	5-11	5-11
108	17-9	17-3	54	5-9	5-10
107	17-6	17-1	53	5-7	5-8
106	17-2	16-10	52	5-5	5-6
105	16-11	16-7	51	5-2	5-4
104	16-7	16-5	50	5-1	5-3
103	16-3	16-1	49	4-11	5-1
102	15-11	15-9	48	4-10	4-11
101	15-9	15-6	47	4-8	4-9
100	15-7	15-3	46	4-7	4-7
99	15-5	14-11	45	4-5	4-6
98	15-3	14-8	44	4-3	4-4
97	15-1	14-5	43	4-1	4-2
96	14-10	14-2	42	4-0	4-0
95	14-6	13-11	41	3-11	3-11
94	14-3	13-8	40	3-10	3-10
93	13-11	13-4	39	3-9	3-9
92	13-7	13-0	38	3-8	3-8
91	13-2	12-10	37	3-7	3-8
90	12-11	12-8	36	3-6	3-7
89	12-9	12-6	35	3-5	3-6
88	12-7	12-4	34	3-4	3-5
87	12-5	12-2	33	3-3	3-4
86	12-3	11-11	32	3-3	3-2
85	12-1	11-7	31	3-2	3-1
84	11-9	11-3	30	3-1	3-0
83	11-4	11-0	29	3-0	3-0
82	11-0	10-10	28	2-11	2-11
81	10-10	10-9	27	2-11	2-10
80	10-8	10-7	26	2-10	2-10
79	10-7	10-5	25	2-9	2-9
78	10-5	10-4	24	2-8	2-8
77	10-4	10-2	23	2-8	2-8
76	10-2	10-1	22	2-7	2-7
75	10-0	9-10	21	2-6	2-6
74	9-8	9-6	20	2-6	2-6
73	9-5	9-4	19	2-5	2-5
72	9-2	9-1	18	2-5	2-5
71	8-11	8-11	17	2-4	2-4
70	8-9	8-9	16	2-4	2-4
69	8-7	8-6	15	2-3	2-3
68	8-5	8-4	14	2-3	2-3
67	8-3	8-2	13	2-2	2-2
66	8-1	8-0	12	2-2	2-2
65	7-10	7-10	11	2-1	2-1
64	7-8	7-7	10	2-1	2-1
63	7-6	7-5	9	2-0	2-0
62	7-3	7-3	8	1-11	1-11
61	7-1	7-0	7	1-11	1-11
60	6-10	6-10	6	1-10	1-10
59	6-8	6-8	5	1-9	1-9
58	6-6	6-6			

TABLE 8

STANDARD SCORE EQUIVALENTS OR VOCABULARY QUOTIENTS (V.Q.) OF RAW SCORES FOR FORM A

Raw Scores	Chronological Ages																		
	17-6 to 18-5	16-6 to 17-5	15-6 to 16-5	14-6 to 15-5	13-6 to 14-5	12-6 to 13-5	11-6 to 12-5	10-6 to 11-5	9-6 to 10-5	8-6 to 9-5	7-6 to 8-5	6-6 to 7-5	5-6 to 6-5	4-9 to 5-5	4-3 to 4-8	3-9 to 4-2	3-3 to 3-8	2-9 to 3-2	2-3 to 2-8
150	141	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
149	140	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
148	139	146	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
147	138	144	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
146	137	143	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
145	136	142	145	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
144	135	141	144	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
143	134	140	143	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
142	133	139	142	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
141	132	138	141	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
140	131	137	140	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
139	130	136	139	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
138	129	135	138	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
137	128	134	137	145	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
136	127	133	135	144	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
135	126	132	134	143	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
134	125	131	133	141	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
133	124	130	132	140	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
132	123	129	131	139	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
131	122	127	130	138	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
130	121	126	129	137	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
129	120	125	128	136	145	-	-	-	-	-	-	-	-	-	-	-	-	-	-
128	119	124	127	135	143	-	-	-	-	-	-	-	-	-	-	-	-	-	-
127	118	123	126	134	142	-	-	-	-	-	-	-	-	-	-	-	-	-	-
126	117	122	125	132	141	145	-	-	-	-	-	-	-	-	-	-	-	-	-
125	116	121	124	131	140	144	-	-	-	-	-	-	-	-	-	-	-	-	-
124	115	120	123	130	138	143	-	-	-	-	-	-	-	-	-	-	-	-	-
123	114	119	122	129	137	142	-	-	-	-	-	-	-	-	-	-	-	-	-
122	113	118	121	128	136	140	-	-	-	-	-	-	-	-	-	-	-	-	-
121	112	117	120	127	135	139	-	-	-	-	-	-	-	-	-	-	-	-	-
120	111	116	119	126	133	138	-	-	-	-	-	-	-	-	-	-	-	-	-
119	110	115	118	125	132	136	-	-	-	-	-	-	-	-	-	-	-	-	-
118	109	114	117	124	131	135	-	-	-	-	-	-	-	-	-	-	-	-	-
117	108	113	116	122	130	134	-	-	-	-	-	-	-	-	-	-	-	-	-
116	107	112	114	121	128	133	-	-	-	-	-	-	-	-	-	-	-	-	-
115	106	110	113	120	127	131	145	-	-	-	-	-	-	-	-	-	-	-	-
114	105	109	112	119	126	130	143	-	-	-	-	-	-	-	-	-	-	-	-
113	104	108	111	118	125	129	142	-	-	-	-	-	-	-	-	-	-	-	-
112	103	107	110	117	123	127	140	145	-	-	-	-	-	-	-	-	-	-	-
111	102	106	109	116	122	126	139	143	-	-	-	-	-	-	-	-	-	-	-
110	101	105	108	115	121	125	137	142	-	-	-	-	-	-	-	-	-	-	-
109	100	104	107	114	120	124	136	140	-	-	-	-	-	-	-	-	-	-	-
108	99	103	106	112	118	122	134	139	-	-	-	-	-	-	-	-	-	-	-
107	98	102	105	111	117	121	133	137	-	-	-	-	-	-	-	-	-	-	-
106	97	101	104	110	116	120	131	136	-	-	-	-	-	-	-	-	-	-	-
105	96	100	103	109	115	118	130	134	146	-	-	-	-	-	-	-	-	-	-
104	95	99	102	108	113	117	128	133	144	-	-	-	-	-	-	-	-	-	-
103	94	98	101	107	112	116	127	131	143	-	-	-	-	-	-	-	-	-	-
102	93	97	100	106	111	115	126	130	141	-	-	-	-	-	-	-	-	-	-
101	92	96	99	105	110	113	124	128	140	-	-	-	-	-	-	-	-	-	-

TABLE 8 (cont'd)

STANDARD SCORE EQUIVALENTS OR VOCABULARY QUOTIENTS (V.Q.) OF RAW SCORES FOR FORM A

	Chronological Ages																		
Raw Scores	17-6 to 18-5	16-6 to 17-5	15-6 to 16-5	14-6 to 15-5	13-6 to 14-5	12-6 to 13-5	11-6 to 12-5	10-6 to 11-5	9-6 to 10-5	8-6 to 9-5	7-6 to 8-5	6-6 to 7-5	5-6 to 6-5	4-9 to 5-5	4-3 to 4-8	3-9 to 4-2	3-3 to 3-8	2-9 to 3-2	2-3 to 2-8
100	91	95	98	103	108	112	123	127	138	-	-	-	-	-	-	-	-	-	-
99	90	94	97	102	107	111	121	125	137	146	-	-	-	-	-	-	-	-	-
98	89	92	96	101	106	109	120	124	135	144	-	-	-	-	-	-	-	-	-
97	88	91	94	100	105	108	118	122	133	143	-	-	-	-	-	-	-	-	-
96	87	90	93	99	103	107	117	121	132	141	-	-	-	-	-	-	-	-	-
95	86	89	92	98	102	106	115	119	130	139	-	-	-	-	-	-	-	-	-
94	85	88	91	97	101	104	114	118	129	138	-	-	-	-	-	-	-	-	-
93	84	87	90	96	100	103	112	116	127	136	-	-	-	-	-	-	-	-	-
92	83	86	89	95	98	102	111	115	126	134	146	-	-	-	-	-	-	-	-
91	82	85	88	93	97	100	109	113	124	133	144	-	-	-	-	-	-	-	-
90	81	84	87	92	96	99	108	112	123	131	142	-	-	-	-	-	-	-	-
89	80	83	86	91	95	98	106	110	121	129	140	-	-	-	-	-	-	-	-
88	79	82	85	90	93	97	105	109	119	128	139	-	-	-	-	-	-	-	-
87	78	81	84	89	92	95	103	107	118	126	137	-	-	-	-	-	-	-	-
86	77	80	83	88	91	94	102	106	116	124	135	-	-	-	-	-	-	-	-
85	76	79	82	87	90	93	101	104	115	122	133	-	-	-	-	-	-	-	-
84	75	78	81	86	88	91	99	103	113	121	132	145	-	-	-	-	-	-	-
83	74	77	80	85	87	90	98	101	112	119	130	143	-	-	-	-	-	-	-
82	73	76	79	83	86	89	96	100	110	118	128	141	-	-	-	-	-	-	-
81	72	74	78	82	85	88	95	98	109	116	126	139	-	-	-	-	-	-	-
80	71	73	77	81	83	86	93	97	107	114	125	137	-	-	-	-	-	-	-
79	70	72	76	80	82	85	92	95	105	113	123	135	-	-	-	-	-	-	-
78	69	71	75	79	81	84	90	94	104	111	121	133	145	-	-	-	-	-	-
77	68	70	73	78	80	82	89	92	102	109	120	131	143	-	-	-	-	-	-
76	67	69	72	77	78	81	87	91	101	108	118	129	141	-	-	-	-	-	-
75	66	68	71	76	77	80	86	89	99	106	116	127	139	145	-	-	-	-	-
74	65	67	70	74	76	79	84	88	98	105	114	125	137	144	-	-	-	-	-
73	64	66	69	73	75	77	83	86	96	103	113	124	135	142	-	145	-	-	-
72	63	65	68	72	73	76	81	85	95	101	111	122	133	140	146	143	-	-	-
71	62	64	67	71	72	75	80	83	93	100	109	120	131	138	144	142	-	-	-
70	61	63	66	70	71	73	79	82	91	98	107	118	129	136	142	141	-	-	-
69	60	62	65	69	70	72	77	80	90	96	106	116	127	134	141	139	-	-	-
68	59	61	64	68	68	71	76	79	88	95	104	114	125	133	139	138	-	-	-
67	58	60	63	67	67	70	74	77	87	93	102	112	123	131	137	136	-	-	-
66	57	58	62	66	66	68	73	76	85	91	100	110	121	129	135	135	-	-	-
65	56	57	61	64	65	67	71	74	84	90	99	108	119	127	134	133	145	-	-
64	55	56	60	63	63	66	70	73	82	88	97	106	117	125	132	132	144	-	-
63	-	55	59	62	62	64	68	71	81	86	95	104	115	123	130	130	142	-	-
62	-	-	58	61	61	63	67	69	79	85	93	102	113	122	129	129	141	-	-
61	-	-	57	60	60	62	65	68	77	83	92	100	111	120	127	127	139	-	-
60	-	-	56	59	58	61	64	66	76	81	90	98	109	118	125	126	138	-	-
59	-	-	55	58	57	59	62	65	74	80	88	96	107	116	123	125	136	-	-
58	-	-	-	57	56	58	61	63	73	78	86	95	105	114	122	123	134	145	-
57	-	-	-	56	55	57	59	62	71	76	85	93	103	112	120	122	133	143	-
56	-	-	-	-	-	55	58	60	70	75	83	91	101	111	118	120	131	141	-
55	-	-	-	-	-	-	56	59	68	73	81	89	99	109	116	119	130	140	-
54	-	-	-	-	-	-	55	57	67	71	79	87	97	107	115	117	128	138	-
53	-	-	-	-	-	-	-	56	65	70	78	85	95	105	113	116	127	137	-
52	-	-	-	-	-	-	-	-	63	68	76	83	93	103	111	114	125	135	-
51	-	-	-	-	-	-	-	-	62	66	74	81	91	101	109	113	124	134	-

TABLE 8 (cont'd)

STANDARD SCORE EQUIVALENTS OR VOCABULARY QUOTIENTS (V.Q.) OF RAW SCORES FOR FORM A

	Chronological Ages																		
aw cores	17-6 to 18-5	16-6 to 17-5	15-6 to 16-5	14-6 to 15-5	13-6 to 14-5	12-6 to 13-5	11-6 to 12-5	10-6 to 11-5	9-6 to 10-5	8-6 to 9-5	7-6 to 8-5	6-6 to 7-5	5-6 to 6-5	4-9 to 5-5	4-3 to 4-8	3-9 to 4-2	3-3 to 3-8	2-9 to 3-2	2-3 to 2-8
50	-	-	-	-	-	-	-	-	60	65	72	79	89	100	108	111	122	132	-
49	-	-	-	-	-	-	-	-	59	63	71	77	87	98	106	110	121	131	-
48	-	-	-	-	-	-	-	-	57	61	69	75	85	96	104	109	119	129	-
47	-	-	-	-	-	-	-	-	56	60	67	73	83	94	102	107	117	128	-
46	-	-	-	-	-	-	-	-	-	58	65	71	81	92	101	106	116	126	145
45	-	-	-	-	-	-	-	-	-	56	64	69	79	90	99	104	114	124	144
44	-	-	-	-	-	-	-	-	-	55	62	68	77	89	97	103	113	123	142
43	-	-	-	-	-	-	-	-	-	-	60	66	75	87	96	101	111	121	140
42	-	-	-	-	-	-	-	-	-	-	58	64	73	85	94	100	110	120	138
41	-	-	-	-	-	-	-	-	-	-	57	62	71	83	92	98	108	118	137
40	-	-	-	-	-	-	-	-	-	-	55	60	69	81	90	97	107	117	135
39	-	-	-	-	-	-	-	-	-	-	-	58	67	79	89	96	105	115	133
38	-	-	-	-	-	-	-	-	-	-	-	56	65	78	87	94	104	114	131
37	-	-	-	-	-	-	-	-	-	-	-	-	63	76	85	93	102	112	129
36	-	-	-	-	-	-	-	-	-	-	-	-	61	74	83	91	101	110	128
35	-	-	-	-	-	-	-	-	-	-	-	-	59	72	82	90	99	109	126
34	-	-	-	-	-	-	-	-	-	-	-	-	57	70	80	88	97	107	124
33	-	-	-	-	-	-	-	-	-	-	-	-	55	68	78	87	96	106	122
32	-	-	-	-	-	-	-	-	-	-	-	-	-	67	76	85	94	104	121
31	-	-	-	-	-	-	-	-	-	-	-	-	-	65	75	84	93	103	119
30	-	-	-	-	-	-	-	-	-	-	-	-	-	63	73	82	91	101	117
29	-	-	-	-	-	-	-	-	-	-	-	-	-	61	71	81	90	100	115
28	-	-	-	-	-	-	-	-	-	-	-	-	-	59	69	80	88	98	114
27	-	-	-	-	-	-	-	-	-	-	-	-	-	57	68	78	87	96	112
26	-	-	-	-	-	-	-	-	-	-	-	-	-	56	66	77	85	95	110
25	-	-	-	-	-	-	-	-	-	-	-	-	-	-	64	75	84	93	108
24	-	-	-	-	-	-	-	-	-	-	-	-	-	-	63	74	82	92	106
23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	61	72	80	90	105
22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	59	71	79	89	103
21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	57	69	77	87	101
20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	56	68	76	86	99
19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	67	74	84	98
18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	65	73	82	96
17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	64	71	81	94
16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	62	70	79	92
15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	61	68	78	90
14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	59	67	76	89
13	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	58	65	75	87
12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	56	63	73	85
11	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	55	62	72	83
10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	60	70	82
9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	59	69	80
8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	57	67	78
7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	56	65	76
6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	64	74
5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	62	73

TABLE 8

STANDARD SCORE EQUIVALENTS OR VOCABULARY QUOTIENTS (V.Q.) OF RAW SCORES FOR FORM B

	Chronological Ages																		
Raw Scores	17-6 to 18-5	16-6 to 17-5	15-6 to 16-5	14-6 to 15-5	13-6 to 14-5	12-6 to 13-5	11-6 to 12-5	10-6 to 11-5	9-6 to 10-5	8-6 to 9-5	7-6 to 8-5	6-6 to 7-5	5-6 to 6-5	4-9 to 5-5	4-3 to 4-8	3-9 to 4-2	3-3 to 3-8	2-9 to 3-2	2-3 to 2-8
150	140	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
149	139	145	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
148	138	144	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
147	137	143	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
146	136	142	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
145	135	141	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
144	134	140	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
143	133	139	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
142	132	138	145	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
141	131	137	144	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
140	130	136	143	145	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
139	129	135	142	144	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
138	128	134	140	143	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
137	127	133	139	142	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
136	126	132	138	141	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
135	125	130	137	139	145	-	-	-	-	-	-	-	-	-	-	-	-	-	-
134	124	129	136	138	144	-	-	-	-	-	-	-	-	-	-	-	-	-	-
133	123	128	135	137	143	-	-	-	-	-	-	-	-	-	-	-	-	-	-
132	122	127	134	136	142	-	-	-	-	-	-	-	-	-	-	-	-	-	-
131	121	126	132	135	140	-	-	-	-	-	-	-	-	-	-	-	-	-	-
130	120	125	131	134	139	145	-	-	-	-	-	-	-	-	-	-	-	-	-
129	119	124	130	133	138	144	-	-	-	-	-	-	-	-	-	-	-	-	-
128	118	123	129	132	137	143	-	-	-	-	-	-	-	-	-	-	-	-	-
127	117	122	128	131	136	141	-	-	-	-	-	-	-	-	-	-	-	-	-
126	116	121	127	130	135	140	-	-	-	-	-	-	-	-	-	-	-	-	-
125	115	120	126	128	134	139	-	-	-	-	-	-	-	-	-	-	-	-	-
124	114	119	124	127	133	138	-	-	-	-	-	-	-	-	-	-	-	-	-
123	113	118	123	126	131	137	-	-	-	-	-	-	-	-	-	-	-	-	-
122	112	117	122	125	130	135	-	-	-	-	-	-	-	-	-	-	-	-	-
121	111	115	121	124	129	134	-	-	-	-	-	-	-	-	-	-	-	-	-
120	110	114	120	123	128	133	-	-	-	-	-	-	-	-	-	-	-	-	-
119	109	113	119	122	127	132	146	-	-	-	-	-	-	-	-	-	-	-	-
118	108	112	118	121	126	131	144	-	-	-	-	-	-	-	-	-	-	-	-
117	107	111	116	120	125	130	143	-	-	-	-	-	-	-	-	-	-	-	-
116	106	110	115	119	124	128	141	145	-	-	-	-	-	-	-	-	-	-	-
115	105	109	114	117	122	127	140	144	-	-	-	-	-	-	-	-	-	-	-
114	103	108	113	116	121	126	139	142	-	-	-	-	-	-	-	-	-	-	-
113	102	107	112	115	120	125	137	141	-	-	-	-	-	-	-	-	-	-	-
112	101	106	111	114	119	124	136	140	-	-	-	-	-	-	-	-	-	-	-
111	100	105	110	113	118	122	134	138	-	-	-	-	-	-	-	-	-	-	-
110	99	104	108	112	117	121	133	137	-	-	-	-	-	-	-	-	-	-	-
109	98	103	107	111	116	120	132	135	-	-	-	-	-	-	-	-	-	-	-
108	97	102	106	110	115	119	130	134	-	-	-	-	-	-	-	-	-	-	-
107	96	100	105	109	113	118	129	133	-	-	-	-	-	-	-	-	-	-	-
106	95	99	104	108	112	117	127	131	-	-	-	-	-	-	-	-	-	-	-
105	94	98	103	106	111	115	126	130	-	-	-	-	-	-	-	-	-	-	-
104	93	97	101	105	110	114	125	129	145	-	-	-	-	-	-	-	-	-	-
103	92	96	100	104	109	113	123	127	143	-	-	-	-	-	-	-	-	-	-
102	91	95	99	103	108	112	122	126	142	-	-	-	-	-	-	-	-	-	-
101	90	94	98	102	107	111	121	124	140	-	-	-	-	-	-	-	-	-	-

TABLE 8 (cont'd)

STANDARD SCORE EQUIVALENTS OR VOCABULARY QUOTIENTS (V.Q.) OF RAW SCORES FOR FORM B

Raw Scores	Chronological Ages 17-6 to 18-5	16-6 to 17-5	15-6 to 16-5	14-6 to 15-5	13-6 to 14-5	12-6 to 13-5	11-6 to 12-5	10-6 to 11-5	9-6 to 10-5	8-6 to 9-5	7-6 to 8-5	6-6 to 7-5	5-6 to 6-5	4-9 to 5-5	4-3 to 4-8	3-9 to 4-2	3-3 to 3-8	2-9 to 3-2	2-3 to 2-8
100	89	93	97	101	106	109	119	123	139	146	-	-	-	-	-	-	-	-	-
99	88	92	96	100	104	108	118	122	137	144	-	-	-	-	-	-	-	-	-
98	87	91	95	99	103	107	116	120	135	143	-	-	-	-	-	-	-	-	-
97	86	90	93	98	102	106	115	119	134	141	-	-	-	-	-	-	-	-	-
96	85	89	92	97	101	105	114	118	132	140	-	-	-	-	-	-	-	-	-
95	84	88	91	95	100	104	112	116	131	138	-	-	-	-	-	-	-	-	-
94	83	86	90	94	99	102	111	115	129	136	-	-	-	-	-	-	-	-	-
93	82	85	89	93	98	101	109	113	127	135	-	-	-	-	-	-	-	-	-
92	81	84	88	92	97	100	108	112	126	133	145	-	-	-	-	-	-	-	-
91	80	83	87	91	95	99	107	111	124	131	143	-	-	-	-	-	-	-	-
90	79	82	85	90	94	98	105	109	122	130	142	-	-	-	-	-	-	-	-
89	78	81	84	89	93	97	104	108	121	128	140	-	-	-	-	-	-	-	-
88	77	80	83	88	92	95	102	107	119	127	138	-	-	-	-	-	-	-	-
87	76	79	82	87	91	94	101	105	118	125	136	-	-	-	-	-	-	-	-
86	75	78	81	86	90	93	100	104	116	123	135	-	-	-	-	-	-	-	-
85	74	77	80	85	89	92	98	102	114	122	133	-	-	-	-	-	-	-	-
84	73	76	79	83	88	91	97	101	113	120	131	146	-	-	-	-	-	-	-
83	72	75	77	82	86	89	96	100	111	119	129	144	-	-	-	-	-	-	-
82	71	74	76	81	85	88	94	98	110	117	128	142	-	-	-	-	-	-	-
81	70	73	75	80	84	87	93	97	108	115	126	140	-	-	-	-	-	-	-
80	69	71	74	79	83	86	91	96	106	114	124	138	-	-	-	-	-	-	-
79	68	70	73	78	82	85	90	94	105	112	123	136	-	-	-	-	-	-	-
78	67	69	72	77	81	84	89	93	103	110	121	134	-	-	-	-	-	-	-
77	66	68	71	76	80	82	87	91	101	109	119	132	-	-	-	-	-	-	-
76	65	67	69	75	79	81	86	90	100	107	117	130	145	-	-	-	-	-	-
75	64	66	68	74	77	80	84	89	98	106	116	128	142	-	-	-	-	-	-
74	63	65	67	72	76	79	83	87	97	104	114	126	140	145	-	-	-	-	-
73	62	64	66	71	75	78	82	86	95	102	112	124	138	144	-	-	-	-	-
72	61	63	65	70	74	76	80	85	93	101	110	122	136	142	-	-	-	-	-
71	60	62	64	69	73	75	79	83	92	99	109	120	134	140	-	-	-	-	-
70	59	61	63	68	72	74	77	82	90	98	107	118	132	138	-	-	-	-	-
69	58	60	61	67	71	73	76	80	89	96	105	116	130	136	146	146	-	-	-
68	57	59	60	66	70	72	75	79	87	94	104	114	127	134	144	144	-	-	-
67	56	58	59	65	68	71	73	78	85	93	102	112	125	132	143	143	-	-	-
66	55	56	58	64	67	69	72	76	84	91	100	110	123	131	141	141	-	-	-
65	-	55	57	63	66	68	70	75	82	89	98	108	121	129	139	139	-	-	-
64	-	-	56	61	65	67	69	74	80	88	97	106	119	127	137	137	146	-	-
63	-	-	55	60	64	66	68	72	79	86	95	104	117	125	135	136	144	-	-
62	-	-	-	59	63	65	66	71	77	85	93	102	115	123	133	134	143	-	-
61	-	-	-	58	62	63	65	69	76	83	92	100	112	121	131	132	141	-	-
60	-	-	-	57	61	62	64	68	74	81	90	99	110	119	129	131	139	-	-
59	-	-	-	56	59	61	62	67	72	80	88	97	108	117	127	129	138	-	-
58	-	-	-	55	58	60	61	65	71	78	86	95	106	116	125	127	136	-	-
57	-	-	-	-	57	59	59	64	69	77	85	93	104	114	123	126	135	-	-
56	-	-	-	-	56	58	58	63	68	75	83	91	102	112	121	124	133	145	-
55	-	-	-	-	55	56	57	61	66	73	81	89	100	110	119	122	131	143	-
54	-	-	-	-	-	55	55	60	64	72	79	87	97	108	117	120	130	141	-
53	-	-	-	-	-	-	-	58	63	70	78	85	95	106	115	119	128	140	-
52	-	-	-	-	-	-	-	57	61	68	76	83	93	104	113	117	127	138	-
51	-	-	-	-	-	-	-	56	59	67	74	81	91	103	112	115	125	136	-

TABLE 8 (cont'd)

STANDARD SCORE EQUIVALENTS OR VOCABULARY QUOTIENTS (V.Q.) OF RAW SCORES FOR FORM B

	Chronological Ages																		
Raw Scores	17-6 to 18-5	16-6 to 17-5	15-6 to 16-5	14-6 to 15-5	13-6 to 14-5	12-6 to 13-5	11-6 to 12-5	10-6 to 11-5	9-6 to 10-5	8-6 to 9-5	7-6 to 8-5	6-6 to 7-5	5-6 to 6-5	4-9 to 5-5	4-3 to 4-8	3-9 to 4-2	3-3 to 3-8	2-9 to 3-2	2-3 to 2-8
50	-	-	-	-	-	-	-	-	58	65	73	79	89	101	110	114	124	135	-
49	-	-	-	-	-	-	-	-	56	64	71	77	87	99	108	112	122	133	-
48	-	-	-	-	-	-	-	-	55	62	69	75	85	97	106	110	120	131	-
47	-	-	-	-	-	-	-	-	-	60	67	73	82	95	104	108	119	130	-
46	-	-	-	-	-	-	-	-	-	59	66	71	80	93	102	107	117	128	145
45	-	-	-	-	-	-	-	-	-	57	64	69	78	91	100	105	116	126	143
44	-	-	-	-	-	-	-	-	-	56	62	67	76	90	98	103	114	124	142
43	-	-	-	-	-	-	-	-	-	-	60	65	74	88	96	102	112	122	140
42	-	-	-	-	-	-	-	-	-	-	59	63	72	86	94	100	111	121	138
41	-	-	-	-	-	-	-	-	-	-	57	61	70	84	92	98	109	119	136
40	-	-	-	-	-	-	-	-	-	-	55	59	67	82	90	97	108	118	135
39	-	-	-	-	-	-	-	-	-	-	-	57	65	80	88	95	106	116	133
38	-	-	-	-	-	-	-	-	-	-	-	55	63	78	86	93	105	114	131
37	-	-	-	-	-	-	-	-	-	-	-	-	61	76	84	91	103	112	129
36	-	-	-	-	-	-	-	-	-	-	-	-	59	75	82	90	101	111	128
35	-	-	-	-	-	-	-	-	-	-	-	-	57	73	81	88	100	109	126
34	-	-	-	-	-	-	-	-	-	-	-	-	55	71	79	86	98	107	124
33	-	-	-	-	-	-	-	-	-	-	-	-	-	69	77	85	97	106	122
32	-	-	-	-	-	-	-	-	-	-	-	-	-	67	75	83	95	104	121
31	-	-	-	-	-	-	-	-	-	-	-	-	-	65	73	81	93	102	119
30	-	-	-	-	-	-	-	-	-	-	-	-	-	63	71	80	92	101	117
29	-	-	-	-	-	-	-	-	-	-	-	-	-	62	69	78	90	99	115
28	-	-	-	-	-	-	-	-	-	-	-	-	-	60	67	76	89	97	114
27	-	-	-	-	-	-	-	-	-	-	-	-	-	58	65	74	87	95	112
26	-	-	-	-	-	-	-	-	-	-	-	-	-	56	63	73	85	94	110
25	-	-	-	-	-	-	-	-	-	-	-	-	-	-	61	71	84	92	108
24	-	-	-	-	-	-	-	-	-	-	-	-	-	-	59	69	82	90	107
23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	57	68	81	89	105
22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	55	66	79	87	103
21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	64	78	85	101
20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	62	76	83	100
19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	61	74	82	98
18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	59	73	80	96
17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	57	71	78	94
16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	56	70	77	93
15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	68	75	91
14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	66	73	89
13	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	65	72	87
12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	63	70	86
11	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	62	68	84
10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	60	66	82
9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	58	65	80
8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	57	63	79
7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	55	61	77
6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	60	75
5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	58	73

TABLE 9
PERCENTILE SCORE EQUIVALENTS (%iles) OF RAW SCORES FOR FORM A

	Chronological Ages																		
Raw Scores	17-6 to 18-5	16-6 to 17-5	15-6 to 16-5	14-6 to 15-5	13-6 to 14-5	12-6 to 13-5	11-6 to 12-5	10-6 to 11-5	9-6 to 10-5	8-6 to 9-5	7-6 to 8-5	6-6 to 7-5	5-6 to 6-5	4-9 to 5-5	4-3 to 4-8	3-9 to 4-2	3-3 to 3-8	2-9 to 3-2	2-3 to 2-8
140	99	99	99	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
139	98	98	98	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
138	97	98	98	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
137	97	97	98	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
136	96	97	98	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
135	94	97	98	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
134	92	96	97	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
133	91	96	97	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
132	91	96	96	99	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
131	90	95	96	98	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
130	89	94	95	98	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
129	88	93	95	98	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
128	88	92	95	97	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
127	87	92	94	97	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
126	86	91	94	97	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
125	85	91	93	97	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
124	84	90	93	97	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
123	82	89	92	96	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
122	81	88	91	96	-	99	-	-	-	-	-	-	-	-	-	-	-	-	-
121	80	87	90	95	-	98	-	-	-	-	-	-	-	-	-	-	-	-	-
120	79	85	89	95	-	98	-	-	-	-	-	-	-	-	-	-	-	-	-
119	77	83	88	94	-	98	-	-	-	-	-	-	-	-	-	-	-	-	-
118	75	82	87	94	99	97	-	-	-	-	-	-	-	-	-	-	-	-	-
117	72	81	85	93	98	97	-	-	-	-	-	-	-	-	-	-	-	-	-
116	69	78	83	91	98	96	-	-	-	-	-	-	-	-	-	-	-	-	-
115	66	76	81	90	97	96	-	-	-	-	-	-	-	-	-	-	-	-	-
114	64	74	79	89	96	96	-	-	-	-	-	-	-	-	-	-	-	-	-
113	62	72	77	89	95	96	-	-	-	-	-	-	-	-	-	-	-	-	-
112	59	69	76	87	95	94	-	-	-	-	-	-	-	-	-	-	-	-	-
111	56	67	73	85	94	94	-	-	-	-	-	-	-	-	-	-	-	-	-
110	53	65	70	83	92	93	99	-	-	-	-	-	-	-	-	-	-	-	-
109	50	62	67	82	91	93	98	-	-	-	-	-	-	-	-	-	-	-	-
108	47	60	65	80	89	92	98	-	-	-	-	-	-	-	-	-	-	-	-
107	44	57	62	78	87	91	98	-	-	-	-	-	-	-	-	-	-	-	-
106	42	54	59	76	85	89	98	-	-	-	-	-	-	-	-	-	-	-	-
105	40	51	56	74	83	87	98	-	-	-	-	-	-	-	-	-	-	-	-
104	37	49	54	72	82	86	96	99	-	-	-	-	-	-	-	-	-	-	-
103	34	47	52	70	80	84	95	98	-	-	-	-	-	-	-	-	-	-	-
102	32	45	50	68	79	83	94	98	99	-	-	-	-	-	-	-	-	-	-
101	31	42	48	65	78	81	93	97	98	-	-	-	-	-	-	-	-	-	-
100	28	40	45	63	77	80	92	96	98	-	-	-	-	-	-	-	-	-	-
99	26	38	42	60	75	79	92	95	98	-	-	-	-	-	-	-	-	-	-
98	23	34	39	58	71	77	90	94	98	-	-	-	-	-	-	-	-	-	-
97	22	32	35	55	67	74	88	92	98	-	-	-	-	-	-	-	-	-	-
96	20	29	33	52	64	72	86	92	97	-	-	-	-	-	-	-	-	-	-
95	18	26	31	49	61	71	85	90	97	99	99	-	-	-	-	-	-	-	-
94	17	22	29	46	57	68	83	88	96	98	98	-	-	-	-	-	-	-	-
93	16	20	27	42	53	64	80	86	95	98	98	-	-	-	-	-	-	-	-
92	15	18	25	38	50	60	78	84	95	98	98	-	-	-	-	-	-	-	-
91	14	16	23	36	46	56	76	82	93	98	98	-	-	-	-	-	-	-	-

TABLE 9 (cont'd)
PERCENTILE SCORE EQUIVALENTS (%iles) OF RAW SCORES FOR FORM A

	Chronological Ages																		
Raw Scores	17-6 to 18-5	16-6 to 17-5	15-6 to 16-5	14-6 to 15-5	13-6 to 14-5	12-6 to 13-5	11-6 to 12-5	10-6 to 11-5	9-6 to 10-5	8-6 to 9-5	7-6 to 8-5	6-6 to 7-5	5-6 to 6-5	4-9 to 5-5	4-3 to 4-8	3-9 to 4-2	3-3 to 3-8	2-9 to 3-2	2-3 to 2-8
90	13	14	21	33	43	52	75	80	92	98	98	-	-	-	-	-	-	-	-
89	11	13	19	30	41	50	73	78	90	97	98	-	-	-	-	-	-	-	-
88	10	12	17	27	39	45	70	75	90	97	98	-	-	-	-	-	-	-	-
87	8	10	15	23	36	40	66	70	88	96	98	-	-	-	-	-	-	-	-
86	5	8	13	20	32	37	62	67	86	95	97	-	-	-	-	-	-	-	-
85	4	6	12	18	28	34	56	64	84	94	96	-	-	-	-	-	-	-	-
84	3	5	11	16	26	31	50	60	82	93	96	-	-	-	-	-	-	-	-
83	3	3	9	14	24	27	45	57	79	91	96	-	-	-	-	-	-	-	-
82	2	3	7	12	21	23	40	51	75	89	96	-	-	-	-	-	-	-	-
81	2	2	7	10	17	19	36	46	71	87	96	-	-	-	-	-	-	-	-
80	2	2	5	9	14	15	30	41	67	84	94	-	-	-	-	-	-	-	-
79	1	1	4	8	11	13	24	37	62	80	94	99	-	-	-	-	-	-	-
78	-	-	3	7	9	10	22	32	57	75	94	98	-	-	-	-	-	-	-
77	-	-	3	5	7	9	19	28	53	71	92	98	-	-	-	-	-	-	-
76	-	-	2	4	6	7	16	24	48	67	89	96	-	-	-	-	-	-	-
75	-	-	2	4	3	6	13	21	45	63	87	93	-	-	-	-	-	-	-
74	-	-	1	3	3	4	12	19	42	60	84	92	-	-	-	-	-	-	-
73	-	-	-	2	2	3	11	17	38	57	82	92	-	-	-	-	-	-	-
72	-	-	-	2	2	3	9	15	34	53	80	92	99	-	-	-	-	-	-
71	-	-	-	1	1	3	8	13	30	50	76	91	98	-	-	-	-	-	-
70	-	-	-	-	-	3	7	12	28	47	70	88	98	-	-	-	-	-	-
69	-	-	-	-	-	2	6	10	26	44	63	84	97	-	-	-	-	-	-
68	-	-	-	-	-	1	5	9	24	39	58	82	96	-	-	-	-	-	-
67	-	-	-	-	-	-	5	8	22	34	55	78	94	-	-	-	-	-	-
66	-	-	-	-	-	-	4	6	20	31	53	74	92	99	-	-	-	-	-
65	-	-	-	-	-	-	3	5	18	28	50	70	90	98	-	-	-	-	-
64	-	-	-	-	-	-	3	4	15	24	46	66	88	97	-	-	-	-	-
63	-	-	-	-	-	-	2	3	12	20	42	62	85	97	-	-	-	-	-
62	-	-	-	-	-	-	2	2	10	16	40	58	83	95	99	99	-	-	-
61	-	-	-	-	-	-	2	2	9	12	34	54	79	93	97	97	-	-	-
60	-	-	-	-	-	-	1	1	7	10	28	49	76	93	96	96	-	-	-
59	-	-	-	-	-	-	-	-	4	8	21	44	73	89	94	96	-	-	-
58	-	-	-	-	-	-	-	-	2	6	16	40	68	84	91	96	-	-	-
57	-	-	-	-	-	-	-	-	1	5	12	35	62	79	90	96	99	-	-
56	-	-	-	-	-	-	-	-	-	4	10	33	54	75	89	96	98	-	-
55	-	-	-	-	-	-	-	-	-	3	10	28	47	71	88	94	97	-	-
54	-	-	-	-	-	-	-	-	-	2	8	23	43	65	87	92	96	-	-
53	-	-	-	-	-	-	-	-	-	1	6	18	37	60	83	91	96	-	-
52	-	-	-	-	-	-	-	-	-	-	5	12	30	56	80	88	95	-	-
51	-	-	-	-	-	-	-	-	-	-	2	8	25	51	78	85	94	-	-
50	-	-	-	-	-	-	-	-	-	-	1	6	24	45	75	82	94	-	-
49	-	-	-	-	-	-	-	-	-	-	-	4	21	38	68	77	92	-	-
48	-	-	-	-	-	-	-	-	-	-	-	2	18	33	61	71	89	-	-
47	-	-	-	-	-	-	-	-	-	-	-	1	16	29	53	66	87	99	-
46	-	-	-	-	-	-	-	-	-	-	-	-	13	27	44	60	84	98	-
45	-	-	-	-	-	-	-	-	-	-	-	-	10	23	38	55	82	94	-
44	-	-	-	-	-	-	-	-	-	-	-	-	8	20	35	51	80	92	-
43	-	-	-	-	-	-	-	-	-	-	-	-	7	16	32	47	76	89	-
42	-	-	-	-	-	-	-	-	-	-	-	-	6	12	28	44	70	87	99
41	-	-	-	-	-	-	-	-	-	-	-	-	4	11	25	42	65	84	98

TABLE 9 (cont'd)
PERCENTILE SCORE EQUIVALENTS (%iles) OF RAW SCORES FOR FORM A

	Chronological Ages																		
Raw Scores	17-6 to 18-5	16-6 to 17-5	15-6 to 16-5	14-6 to 15-5	13-6 to 14-5	12-6 to 13-5	11-6 to 12-5	10-6 to 11-5	9-6 to 10-5	8-6 to 9-5	7-6 to 8-5	6-6 to 7-5	5-6 to 6-5	4-9 to 5-5	4-3 to 4-8	3-9 to 4-2	3-3 to 3-8	2-9 to 3-2	2-3 to 2-8
40	-	-	-	-	-	-	-	-	-	-	-	-	3	10	21	38	63	82	98
39	-	-	-	-	-	-	-	-	-	-	-	-	3	10	18	35	60	82	96
38	-	-	-	-	-	-	-	-	-	-	-	-	3	9	16	32	56	80	95
37	-	-	-	-	-	-	-	-	-	-	-	-	1	8	14	30	53	76	94
36	-	-	-	-	-	-	-	-	-	-	-	-	-	6	12	28	51	72	93
35	-	-	-	-	-	-	-	-	-	-	-	-	-	5	12	26	49	66	93
34	-	-	-	-	-	-	-	-	-	-	-	-	-	4	11	23	47	63	92
33	-	-	-	-	-	-	-	-	-	-	-	-	-	3	11	20	44	60	90
32	-	-	-	-	-	-	-	-	-	-	-	-	-	2	11	17	41	57	89
31	-	-	-	-	-	-	-	-	-	-	-	-	-	2	10	15	38	56	88
30	-	-	-	-	-	-	-	-	-	-	-	-	-	2	8	12	32	55	87
29	-	-	-	-	-	-	-	-	-	-	-	-	-	2	6	10	29	53	86
28	-	-	-	-	-	-	-	-	-	-	-	-	-	2	4	8	26	52	86
27	-	-	-	-	-	-	-	-	-	-	-	-	-	2	2	7	24	50	84
26	-	-	-	-	-	-	-	-	-	-	-	-	-	1	2	7	21	45	82
25	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	5	18	39	80
24	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	4	13	35	74
23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4	6	33	67
22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4	5	28	63
21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4	4	21	58
20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	2	17	54
19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	2	15	52
18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	2	13	50
17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	9	46
16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	7	38
15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6	28
14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4	23
13	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	17
12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	10
11	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	5
10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4
9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4
8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3
7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2
6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1

TABLE 9
PERCENTILE SCORE EQUIVALENTS (%iles) OF RAW SCORES FOR FORM B

	Chronological Ages																		
Raw Scores	17-6 to 18-5	16-6 to 17-5	15-6 to 16-5	14-6 to 15-5	13-6 to 14-5	12-6 to 13-5	11-6 to 12-5	10-6 to 11-5	9-6 to 10-5	8-6 to 9-5	7-6 to 8-5	6-6 to 7-5	5-6 to 6-5	4-9 to 5-5	4-3 to 4-8	3-9 to 4-2	3-3 to 3-8	2-9 to 3-2	2-3 to 2-8
140	99	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
139	98	99	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
138	97	98	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
137	96	98	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
136	94	98	99	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
135	93	97	98	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
134	92	97	98	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
133	92	97	97	99	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
132	91	96	97	98	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
131	89	95	97	98	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
130	88	95	96	98	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
129	87	94	95	98	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
128	86	93	95	98	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
127	85	91	95	97	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
126	83	90	94	97	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
125	82	90	94	96	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
124	81	88	93	95	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
123	79	86	93	94	-	99	-	-	-	-	-	-	-	-	-	-	-	-	-
122	78	85	93	93	-	98	-	-	-	-	-	-	-	-	-	-	-	-	-
121	76	84	92	93	-	98	-	-	-	-	-	-	-	-	-	-	-	-	-
120	74	82	91	93	-	98	-	-	-	-	-	-	-	-	-	-	-	-	-
119	72	81	90	93	99	98	-	-	-	-	-	-	-	-	-	-	-	-	-
118	69	78	88	93	98	98	-	-	-	-	-	-	-	-	-	-	-	-	-
117	66	76	88	91	98	97	-	-	-	-	-	-	-	-	-	-	-	-	-
116	63	74	86	89	98	96	-	-	-	-	-	-	-	-	-	-	-	-	-
115	61	73	84	88	97	96	99	-	-	-	-	-	-	-	-	-	-	-	-
114	59	72	82	87	96	95	98	-	-	-	-	-	-	-	-	-	-	-	-
113	57	70	81	86	94	94	98	-	-	-	-	-	-	-	-	-	-	-	-
112	54	68	78	84	93	94	98	-	-	-	-	-	-	-	-	-	-	-	-
111	51	65	75	83	91	93	98	-	-	-	-	-	-	-	-	-	-	-	-
110	49	61	73	82	89	92	97	99	-	-	-	-	-	-	-	-	-	-	-
109	46	57	70	81	88	91	97	98	-	-	-	-	-	-	-	-	-	-	-
108	45	54	67	79	86	90	97	98	-	-	-	-	-	-	-	-	-	-	-
107	43	52	63	76	83	89	96	97	-	-	-	-	-	-	-	-	-	-	-
106	41	50	58	73	81	87	96	97	-	-	-	-	-	-	-	-	-	-	-
105	37	46	54	70	79	84	95	96	99	-	-	-	-	-	-	-	-	-	-
104	34	43	51	66	76	82	94	95	98	-	-	-	-	-	-	-	-	-	-
103	31	39	47	62	73	79	93	95	98	-	-	-	-	-	-	-	-	-	-
102	27	35	44	59	69	76	91	94	97	-	-	-	-	-	-	-	-	-	-
101	24	32	41	54	66	73	90	93	97	-	-	-	-	-	-	-	-	-	-
100	23	30	38	50	64	71	89	93	97	-	-	-	-	-	-	-	-	-	-
99	22	28	36	47	62	69	88	92	97	-	-	-	-	-	-	-	-	-	-
98	22	27	33	44	59	67	86	91	96	-	-	-	-	-	-	-	-	-	-
97	21	25	31	41	56	65	83	89	96	-	-	-	-	-	-	-	-	-	-
96	18	23	29	39	54	63	81	88	96	-	-	-	-	-	-	-	-	-	-
95	16	22	27	36	52	61	78	86	95	99	-	-	-	-	-	-	-	-	-
94	15	20	25	34	48	58	77	83	95	98	-	-	-	-	-	-	-	-	-
93	13	18	24	33	46	56	75	82	95	97	-	-	-	-	-	-	-	-	-
92	11	17	23	31	44	54	73	81	95	97	-	-	-	-	-	-	-	-	-
91	11	15	22	30	42	50	70	79	94	97	-	-	-	-	-	-	-	-	-

TABLE 9 (cont'd)

PERCENTILE SCORE EQUIVALENTS (%iles) OF RAW SCORES FOR FORM B

	Chronological Ages																		
Raw Scores	17-6 to 18-5	16-6 to 17-5	15-6 to 16-5	14-6 to 15-5	13-6 to 14-5	12-6 to 13-5	11-6 to 12-5	10-6 to 11-5	9-6 to 10-5	8-6 to 9-5	7-6 to 8-5	6-6 to 7-5	5-6 to 6-5	4-9 to 5-5	4-3 to 4-8	3-9 to 4-2	3-3 to 3-8	2-9 to 3-2	2-3 to 2-8
90	9	13	19	28	41	48	69	77	93	96	-	-	-	-	-	-	-	-	-
89	7	11	17	26	40	47	67	76	92	96	-	-	-	-	-	-	-	-	-
88	6	10	14	25	37	44	64	72	91	96	99	-	-	-	-	-	-	-	-
87	5	9	12	22	33	41	61	69	89	95	98	-	-	-	-	-	-	-	-
86	4	8	11	19	30	37	58	64	87	94	98	-	-	-	-	-	-	-	-
85	3	6	9	16	25	33	55	60	85	93	98	-	-	-	-	-	-	-	-
84	2	5	7	15	21	30	49	56	82	92	97	-	-	-	-	-	-	-	-
83	2	4	6	13	18	26	43	53	80	91	97	-	-	-	-	-	-	-	-
82	1	3	6	11	16	22	37	48	75	89	97	-	-	-	-	-	-	-	-
81	-	3	4	9	14	18	31	43	71	86	96	-	-	-	-	-	-	-	-
80	-	3	4	7	12	15	27	38	67	83	95	99	-	-	-	-	-	-	-
79	-	2	3	6	10	13	22	35	63	79	94	98	-	-	-	-	-	-	-
78	-	2	3	5	9	11	18	30	58	77	92	98	-	-	-	-	-	-	-
77	-	1	2	4	8	10	14	26	51	75	90	97	-	-	-	-	-	-	-
76	-	-	2	4	7	8	12	23	47	72	88	97	-	-	-	-	-	-	-
75	-	-	1	3	5	7	11	20	43	68	84	96	-	-	-	-	-	-	-
74	-	-	-	2	5	6	9	18	38	64	82	94	-	-	-	-	-	-	-
73	-	-	-	2	4	5	7	16	34	59	78	93	-	-	-	-	-	-	-
72	-	-	-	1	3	4	6	15	30	53	76	92	-	-	-	-	-	-	-
71	-	-	-	-	2	4	5	13	28	50	72	91	99	-	-	-	-	-	-
70	-	-	-	-	2	3	4	11	27	46	66	90	98	-	-	-	-	-	-
69	-	-	-	-	2	2	4	10	25	42	61	86	98	-	-	-	-	-	-
68	-	-	-	-	1	2	3	9	23	37	58	80	97	-	-	-	-	-	-
67	-	-	-	-	-	1	3	8	20	32	56	78	96	-	-	-	-	-	-
66	-	-	-	-	-	-	3	6	16	25	52	74	94	99	-	-	-	-	-
65	-	-	-	-	-	-	3	4	13	24	47	69	93	98	-	-	-	-	-
64	-	-	-	-	-	-	2	3	10	20	44	67	91	96	-	-	-	-	-
63	-	-	-	-	-	-	2	2	7	18	40	64	89	95	-	99	-	-	-
62	-	-	-	-	-	-	2	2	5	15	36	59	87	95	-	98	-	-	-
61	-	-	-	-	-	-	1	1	4	12	34	56	84	93	99	98	-	-	-
60	-	-	-	-	-	-	-	-	3	10	30	52	81	91	98	98	-	-	-
59	-	-	-	-	-	-	-	-	2	8	24	46	76	87	98	98	-	-	-
58	-	-	-	-	-	-	-	-	1	6	18	40	70	85	96	98	-	-	-
57	-	-	-	-	-	-	-	-	-	4	16	34	62	81	93	98	-	-	-
56	-	-	-	-	-	-	-	-	-	2	11	30	54	78	92	96	-	-	-
55	-	-	-	-	-	-	-	-	-	1	8	25	48	76	91	95	-	-	-
54	-	-	-	-	-	-	-	-	-	-	6	20	41	73	89	92	99	-	-
53	-	-	-	-	-	-	-	-	-	-	4	16	35	69	86	90	98	-	-
52	-	-	-	-	-	-	-	-	-	-	4	13	30	62	84	90	97	99	-
51	-	-	-	-	-	-	-	-	-	-	3	10	25	56	81	88	96	98	-
50	-	-	-	-	-	-	-	-	-	-	2	7	20	51	77	84	94	98	-
49	-	-	-	-	-	-	-	-	-	-	2	3	16	46	74	80	92	98	-
48	-	-	-	-	-	-	-	-	-	-	2	2	13	41	67	76	90	98	-
47	-	-	-	-	-	-	-	-	-	-	1	1	10	36	59	71	88	98	99
46	-	-	-	-	-	-	-	-	-	-	-	-	8	30	54	69	86	98	98
45	-	-	-	-	-	-	-	-	-	-	-	-	6	24	48	65	84	97	98
44	-	-	-	-	-	-	-	-	-	-	-	-	5	17	42	59	82	97	98
43	-	-	-	-	-	-	-	-	-	-	-	-	4	15	35	55	79	95	98
42	-	-	-	-	-	-	-	-	-	-	-	-	3	13	27	51	73	92	97
41	-	-	-	-	-	-	-	-	-	-	-	-	3	11	24	43	67	90	97

TABLE 9 (cont'd)
PERCENTILE SCORE EQUIVALENTS (%iles) OF RAW SCORES FOR FORM B

	Chronological Ages																		
aw cores	17-6 to 18-5	16-6 to 17-5	15-6 to 16-5	14-6 to 15-5	13-6 to 14-5	12-6 to 13-5	11-6 to 12-5	10-6 to 11-5	9-6 to 10-5	8-6 to 9-5	7-6 to 8-5	6-6 to 7-5	5-6 to 6-5	4-9 to 5-5	4-3 to 4-8	3-9 to 4-2	3-3 to 3-8	2-9 to 3-2	2-3 to 2-8
40	-	-	-	-	-	-	-	-	-	-	-	-	3	9	20	37	64	87	97
39	-	-	-	-	-	-	-	-	-	-	-	-	3	9	18	30	63	85	96
38	-	-	-	-	-	-	-	-	-	-	-	-	2	8	16	23	61	84	95
37	-	-	-	-	-	-	-	-	-	-	-	-	2	6	14	21	57	82	94
36	-	-	-	-	-	-	-	-	-	-	-	-	1	4	11	19	53	75	93
35	-	-	-	-	-	-	-	-	-	-	-	-	-	4	9	16	51	69	92
34	-	-	-	-	-	-	-	-	-	-	-	-	-	4	7	14	47	66	92
33	-	-	-	-	-	-	-	-	-	-	-	-	-	3	7	12	41	62	91
32	-	-	-	-	-	-	-	-	-	-	-	-	-	3	6	12	36	58	89
31	-	-	-	-	-	-	-	-	-	-	-	-	-	3	5	11	32	54	88
30	-	-	-	-	-	-	-	-	-	-	-	-	-	2	4	10	29	50	88
29	-	-	-	-	-	-	-	-	-	-	-	-	-	2	3	9	27	47	88
28	-	-	-	-	-	-	-	-	-	-	-	-	-	2	3	7	25	45	87
27	-	-	-	-	-	-	-	-	-	-	-	-	-	2	2	7	24	41	86
26	-	-	-	-	-	-	-	-	-	-	-	-	-	2	2	6	22	38	84
25	-	-	-	-	-	-	-	-	-	-	-	-	-	2	2	5	18	36	82
24	-	-	-	-	-	-	-	-	-	-	-	-	-	1	2	5	14	32	77
23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	4	12	27	71
22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	10	22	65
21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	7	16	60
20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	5	11	57
19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	4	9	50
18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	3	7	43
17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	6	39
16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	5	34
15	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4	29
14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4	24
13	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	19
12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	16
11	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	11
10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6
9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3
8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2
7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2
6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1

APPENDIX A

PEABODY PICTURE VOCABULARY TEST

Individual Test Record

Name ________________ (Last) ________________ (First) ________________ (Initial) Sex: M F (circle) Grade ________

School ________________ (or address) Teacher ________________ (or parent or phone)

Calculation		DERIVED SCORES			Year	Month	Day
Ceiling item	________	Age equivalent (V.A.)	________	Date	____	____	____
Errors	________	Standard score (V.Q.)	________	Born	____	____	____
Raw score	________	Percentile (%ile)	________	Age	____	____	

Examiner ________________ Time ________________ Code ________________

• • • J-1 • • • [illegible]-2 • • • M-3 • • • A-4 • • • M-5 • • • J-6 • • • J-7 • • • A-8 • • • S-9 • • • O-10 • • • N-11 • • • D-12 • • •

TEST BEHAVIOR

Examples needed:	____ only 1	____ 2 or 3	____ over 3
Type of response:	____ Subject pointed	____ S. called numbers	____ Examiner pointed
Rapport:	____ easily attained	____ *slowly attained*	____ poor rapport
Guessing:	____ prone to guess	____ guessed when asked	____ resisted guessing
Speed of response:	____ fast	____ average	____ slow
Verbalization:	____ talkative	____ average	____ taciturn
Attention span:	____ distractible	____ average	____ very attentive
Perseveration:	____ none noted	____ some	____ frequent
Need for praise:	____ little needed	____ some needed	____ much needed
Other test behavior:			

PHYSICAL CHARACTERISTICS

Motor activity:	____ hyperactive	____ average	____ hypoactive
Sedation:	____ none	____ slight	____ heavy
Ambulation:	____ normal	____ walks with support	____ none
Speech:	____ intelligible	____ fairly intelligible	____ unintelligible
Hearing: necessity to repeat stimulus words	____ never	____ seldom	____ often
	____ S. wore hearing aid	____ S. watched examiner's lips and face closely.	
Vision: distance of eyes from page	____ under 8"	____ average (8"-20")	____ over 20"
	____ S. wore glasses	____ S. owned but did not wear glasses during test.	
Other physical characteristics:			

OTHER INFORMATION (previous tests, dates, scores etc.; teacher estimates of vocabulary, intelligence, achievement; school or work record)

All rights reserved. This form may not be reproduced.

Name____ **Form A**

Item	Resp.	Key	Word	Item	Resp.	Key	Word	Item	Resp.	Key	Word
1		(4)	car	51		(4)	submarine	101		(3)	graduated
2		(3)	cow	52		(4)	thermos	102		(2)	hieroglyphic
3		(1)	baby	53		(3)	projector	103		(1)	orate
4		(2)	girl	54		(4)	group	104		(3)	cascade
5		(1)	ball	55		(3)	tackling	105		(4)	illumination
6		(3)	block	56		(1)	transportation	106		(1)	nape
7		(2)	clown	57		(1)	counter	107		(2)	geneologist
8		(1)	key	58		(2)	ceremony	108		(2)	embossed
9		(4)	can	59		(3)	pod	109		(4)	mercantile
10		(2)	chicken	60		(4)	bronco	110		(2)	encumbered
11		(4)	blowing	61		(3)	directing	111		(4)	entice
12		(2)	fan	62		(4)	funnel	112		(3)	concentric
13		(1)	digging	63		(2)	delight	113		(3)	vitreous
14		(1)	skirt	64		(3)	lecturer	114		(1)	sibling
15		(4)	catching	65		(2)	communication	115		(2)	machete
16		(1)	drum	66		(4)	archer	116		(4)	waif
17		(3)	leaf	67		(1)	stadium	117		(1)	cornice
18		(4)	tying	68		(1)	excavate	118		(3)	timorous
19		(1)	fence	69		(4)	assaulting	119		(1)	fettered
20		(2)	bat	70		(1)	stunt	120		(2)	tartan
21		(4)	bee	71		(1)	meringue	121		(3)	sulky
22		(3)	bush	72		(3)	appliance	122		(4)	obelisk
23		(1)	pouring	73		(4)	chemist	123		(2)	ellipse
24		(1)	sewing	74		(3)	arctic	124		(2)	entomology
25		(4)	wiener	75		(4)	destruction	125		(4)	bumptious
26		(2)	teacher	76		(3)	porter	126		(2)	dormer
27		(3)	building	77		(2)	coast	127		(2)	coniferous
28		(3)	arrow	78		(4)	hoisting	128		(4)	consternation
29		(2)	kangaroo	79		(1)	wailing	129		(3)	obese
30		(3)	accident	80		(2)	coil	130		(4)	gauntlet
31		(3)	nest	81		(3)	kayak	131		(1)	inclement
32		(4)	caboose	82		(2)	sentry	132		(1)	cupola
33		(1)	envelope	83		(4)	furrow	133		(2)	obliterate
34		(2)	picking	84		(1)	beam	134		(3)	burnishing
35		(1)	badge	85		(3)	fragment	135		(1)	bovine
36		(3)	goggles	86		(2)	hovering	136		(4)	eminence
37		(2)	peacock	87		(3)	bereavement	137		(3)	legume
38		(3)	queen	88		(4)	crag	138		(4)	senile
39		(4)	coach	89		(2)	tantrum	139		(2)	deleterious
40		(1)	whip	90		(1)	submerge	140		(4)	raze
41		(4)	net	91		(3)	descend	141		(2)	ambulation
42		(4)	freckle	92		(2)	hassock	142		(1)	cravat
43		(3)	eagle	93		(1)	canine	143		(2)	impale
44		(2)	twist	94		(1)	probing	144		(4)	marsupial
45		(4)	shining	95		(1)	angling	145		(3)	predatory
46		(2)	dial	96		(3)	appraising	146		(1)	incertitude
47		(2)	yawning	97		(4)	confining	147		(2)	imbibe
48		(2)	tumble	98		(4)	precipitation	148		(3)	homunculus
49		(1)	signal	99		(1)	gable	149		(4)	cryptogam
50		(1)	capsule	100		(1)	amphibian	150		(3)	pensile

APPENDIX B

PEABODY PICTURE VOCABULARY TEST

Individual Test Record

Name ____ Sex: M F Grade __________
(Last) (First) (Initial) (circle)

School ______________________________ Teacher ______________________________
(or address) (or parent or phone)

Calculation			DERIVED SCORES			Year	Month	Day
Ceiling item	________	Age equivalent (V.A.)	________		Date	______	______	______
Errors	________	Standard score (V.Q.)	________		Born	______	______	______
Raw score	________	Percentile (%ile)	________		Age	______	______	

Examiner ______________________________ Time ______________ Code ______________

• • • J-1 • • • F-2 • • • M-3 • • • A-4 • • • M-5 • • • J-6 • • • J-7 • • • A-8 • • • S-9 • • • O-10 • • • N-11 • • • D-12 • • •

TEST BEHAVIOR

Examples needed:	____only 1	____2 or 3	____over 3
Type of response:	____Subject pointed	____S. called numbers	____Examiner pointed
Rapport:	____easily attained	____slowly attained	____poor rapport
Guessing:	____prone to guess	____guessed when asked	____resisted guessing
Speed of response:	____fast	____average	____slow
Verbalization:	____talkative	____average	____taciturn
Attention span:	____distractible	____average	____very attentive
Perseveration:	____none noted	____some	____frequent
Need for praise:	____little needed	____some needed	____much needed
Other test behavior:			

PHYSICAL CHARACTERISTICS

Motor activity:	____hyperactive	____average	____hypoactive
Sedation:	____none	____slight	____heavy
Ambulation:	____normal	____walks with support	____none
Speech:	____intelligible	____fairly intelligible	____unintelligible
Hearing: necessity to repeat stimulus words	____never	____seldom	____often
	____S. wore hearing aid	____S. watched examiner's lips and face closely.	
Vision: distance of eyes from page	____under 8"	____average (8"-20")	____over 20"
	____S. wore glasses	____S. owned but did not wear glasses during test.	
Other physical characteristics:			

OTHER INFORMATION

(previous tests, dates, scores etc.; teacher estimates of vocabulary, intelligence, achievement; school or work record)

All rights reserved. This form may not be reproduced.

Name____

Form B

Item	Resp.	Key	Word
1		(2)	table
2		(4)	bus
3		(2)	horse
4		(3)	dog
5		(4)	shoe
6		(4)	finger
7		(3)	boat
8		(2)	children
9		(1)	bell
10		(4)	turtle
11		(2)	climbing
12		(1)	lamp
13		(3)	sitting
14		(2)	jacket
15		(1)	pulling
16		(2)	ring
17		(1)	nail
18		(2)	hitting
19		(3)	tire
20		(3)	ladder
21		(1)	snake
22		(1)	river
23		(4)	ringing
24		(4)	baking
25		(2)	cone
26		(3)	engineer
27		(4)	peeking
28		(1)	kite
29		(1)	rat
30		(1)	time
31		(4)	sail
32		(2)	ambulance
33		(2)	trunk
34		(4)	skiing
35		(2)	hook
36		(1)	tweezers
37		(3)	wasp
38		(2)	barber
39		(3)	parachute
40		(4)	saddle
41		(3)	temperature
42		(1)	captain
43		(2)	whale
44		(4)	cash
45		(1)	balancing
46		(3)	cobweb
47		(3)	pledging
48		(1)	argument
49		(3)	hydrant
50		(4)	binocular
51		(1)	locomotive
52		(2)	hive
53		(4)	reel
54		(1)	insect
55		(1)	gnawing
56		(2)	weapon
57		(3)	bannister
58		(1)	idol
59		(1)	globe
60		(3)	walrus
61		(1)	filing
62		(3)	shears
63		(1)	horror
64		(4)	chef
65		(4)	harvesting
66		(3)	construction
67		(4)	observatory
68		(4)	assistance
69		(2)	erecting
70		(3)	thoroughbred
71		(2)	casserole
72		(4)	ornament
73		(3)	cobbler
74		(2)	autumn
75		(3)	dissatisfaction
76		(4)	scholar
77		(1)	oasis
78		(3)	soldering
79		(3)	astonishment
80		(1)	tread
81		(2)	thatched
82		(1)	jurisprudence
83		(2)	sapling
84		(3)	arch
85		(4)	dwelling
86		(1)	lubricating
87		(2)	pedestrian
88		(3)	vale
89		(3)	jubilant
90		(2)	laden
91		(2)	pursuit
92		(4)	goblet
93		(2)	rodent
94		(3)	confiding
95		(4)	reclining
96		(1)	frisking
97		(2)	moat
98		(3)	salutation
99		(2)	barrier
100		(3)	foal
101		(4)	incandescent
102		(3)	cornucopia
103		(2)	ascending
104		(1)	summit
105		(3)	caster
106		(2)	lobe
107		(3)	patriarch
108		(3)	sampler
109		(3)	ingenious
110		(1)	repose
111		(3)	constrain
112		(1)	tangent
113		(4)	sconce
114		(4)	hoary
115		(1)	pendant
116		(1)	prodigy
117		(2)	casement
118		(1)	quiescent
119		(4)	talon
120		(1)	chevron
121		(4)	feline
122		(2)	cairn
123		(4)	convergence
124		(3)	apothecary
125		(2)	indigent
126		(4)	edifice
127		(3)	scallion
128		(1)	infirm
129		(1)	emaciate
130		(2)	catapult
131		(2)	arable
132		(4)	orifice
133		(3)	renovate
134		(1)	precarious
135		(2)	dromedary
136		(1)	pedagogue
137		(1)	sepal
138		(3)	lethargic
139		(4)	delectation
140		(3)	embellish
141		(1)	osculation
142		(2)	cincture
143		(3)	barrister
144		(3)	carrion
145		(2)	lanate
146		(4)	chirography
147		(1)	mendicant
148		(1)	saltation
149		(2)	florescence
150		(4)	culver

APPENDIX C

COMPARABILITY OF PEABODY PICTURE VOCABULARY TEST SCORES UNDER GROUP AND INDIVIDUAL ADMINISTRATION[1]

Raymond C. Norris, John V. Hottel, and Sadye Brooks
George Peabody College for Teachers

The present study was conducted in connection with the development and standardization of the Peabody Picture Vocabulary Test (PPVT). This test consists of one hundred fifty plates containing four pictures each. The pictures on any one plate all represent objects, or activities, or states of being. In administering the test the examiner provides a stimulus word orally as in "Show me ball". The subject then points to, or in some other way indicates, the picture on the plate which best illustrates the meaning of the stimulus word. The plates are arranged in ascending order of empirically-determined difficulty, and the subject is tested from a basal of eight consecutive correct responses to a ceiling of six errors on eight consecutive plates. The raw score for a subject is his basal score plus the number of items answered correctly up to his ceiling. A more complete description of the instrument has been reported by Dunn (1) in his manual.

Purposes

While the PPVT was designed to be used as an individual examination, the test author and his staff found it advantageous in trying out possible test items on some levels to administer the items to groups rather than individuals. Photographic slides of tentative item plates were projected onto a screen and classroom groups of subjects marked their responses on separate answer sheets. The procedure worked so well that consideration was given to the possibility of collecting the normative data for normal children of late elementary school age in the same manner. However, there was no evidence that group testing would yield scores comparable to those obtained under individual administration.

The primary purpose of the present study was to test the hypothesis that PPVT scores obtained under individual and group administrations would not differ significantly or appreciably. If this hypothesis was upheld, it was believed that the norm sample could be increased greatly while effecting considerable savings in time and expense.

---

[1] This research reported herein was performed pursuant to Contract SAE-6425, United States Office of Education, Department of Health, Education, and Welfare.

The generous cooperation of Lloyd M. Dunn, Project Director, and his project staff, is gratefully acknowledged.

Two other purposes helped determine the design used in the current study. In the preparation of the PPVT, item plates were selected in such a way that two items of nearly comparable difficulty could be obtained from each page. These items were then assigned to forms A and B in such a way as to equate difficulty of the forms for each year level. The authors were interested in finding out how successful this equating process had been and how comparable scores on the two forms might be considered.

The third purpose for the present study was the determination of how much practice effect could be expected on a second administration of the test in alternate form. In assessing the outcomes of special training programs it seemed desirable that practice effect either be eliminated entirely or estimated very closely. In clinical evaluation the second form might be used for verification of diagnosis, and knowledge of how much score discrepancy might be expected due to practice alone would be helpful.

The three major hypotheses to be tested then were:

H_1: PPVT scores obtained under group and individual administration do not differ significantly. $(\mu_G=\mu_I)$

H_2: PPVT scores based on Form A do not differ significantly from those based on Form B. $(\mu_A=\mu_B)$

H_3: PPVT scores will not be significantly altered by the subjects' having had another form of the test earlier. $(\mu_1=\mu_2)$

Experimental Design

The experimental design used in this study is called a Type V Mixed Design by Lindquist (2). With two forms of the test, two types of administration, and two possible orders of administration this design made it possible for the effects of all three treatments to be evaluated as within-subject comparisons. Any given subject received both form A and form B, both an individual and a group administration, and both a first and second administration. The increased precision obtained by using within-subjects rather than a combination between-subjects and within-subjects variability as an error term made it more likely that any differential effects of forms, treatments, or orders would be identified.

The experimental design might be diagrammed as follows:

1st Administration

Form		
A	Group 1	Group 3
B	Group 2	Group 4

Individual Group
Administration

2nd Administration

Form		
A	Group 4	Group 2
B	Group 3	Group 1

Individual Group
Administration

Subjects

The four groups of subjects called for in the experimental design were selected from the fifth grade of one of the Nashville City Schools.[2] Although all members of the four classes were tested, the data were analyzed only for children of appropriate age for grade placement (117 to 137 months in the second month of the school year) and of approximately normal IQ (Kuhlman-Finch IQ scores between 80 and 121). Since it was not possible to constitute random groups for the purposes of this experiment, the selected subjects within each class became an experimental group. While this procedure was not considered ideal, the authors were not able to identify relevant variables such as age, sex, or IQ on which the groups differed significantly. Assignment of students to the classroom groups had been on an essentially random basis at the beginning of the year.

Preliminary screening of subjects on the basis of age-grade placement and intelligence left unequal number of subjects within the four subgroups. To facilitate the analysis subjects were eliminated at random until each group contained the same number as the smallest subgroup. In this way the number of cases per group was reduced to fifteen. Table I provides a summary of sex, IQ, and CA data on the subjects in each experimental group.

TABLE I

BACKGROUND DATA ON EXPERIMENTAL GROUPS

	IQ		CA (Months)		Sex	
	X̄	s	X̄	s	M	F
1	97.47	11.05	124.93	4.33	6	9
2	97.07	10.86	128.53	6.05	9	6
3	101.73	9.91	125.73	4.98	7	8
4	97.73	10.35	126.13	5.01	7	8
Combined	98.50	10.19	126.33	5.18	29	31

Procedure

On the first day of the experiment six examiners administered the appropriate form of the test to individual subjects in the first two groups. The same six examiners also administered the appropriate form of the test to the remaining two intact groups. On the third day of the experiment the same six examiners administered the alternate form of the test to those individuals who had been tested in groups earlier. The subjects who were tested individually on the first day were then given the alternate form in a group presentation. Individual tests were administered to these children in the main room of the school library, a small work-room which adjoins the library, the school clinic, or the private study of the school principal. Each child was brought into a testing room and given the following instructions:

---

[2] The cooperation of Mr. W. H. Oliver, Superintendent of Schools Mr. Rupert Klaus, School Psychologist, along with the principal and teachers involved in the study, is gratefully acknowledged.

> "I am going to show you some pictures. There are four pictures on each page - one, two, three, four (pointing at the same time to an appropriate picture on plate No. Example A). I will say a word, and I want you to tell me the number of the picture that best tells the meaning of the word."

The child was then presented with the three sample plates, Examples A, B, and C, and the proper form of response was established. The examiner then said:

> "That's fine. Now I'm going to show you some more pictures. I'll say a word, and I want you to give me the number of the picture that best tells the meaning of the word: one, two, three, or four. As we get further along, you'll find that you may not be sure of the meaning of the word. I want you to look carefully at all four pictures anyway, and make the best choice you can. Be sure to look carefully at all of the pictures."

As the test progressed, each child was given verbal encouragement, and admonished several times to, "look at all the pictures." Testing was begun at plate No. 70 for each child. The plates were presented in ascending order. If a child committed an error before he had achieved eight successful identifications, the examiner turned back to plate No. 69, and the plates were presented in descending order until a total of eight consecutive pictures were identified. Eight consecutive successful identifications constituted a basal point, and the child was assumed to be able to identify correctly any picture below that point. Once the basal point was established, testing was resumed at the point where the error was committed and proceeded until the child committed six errors out of any eight consecutive presentations (chance performance). At this point it was assumed that he would continue to perform at only chance level.

The number of the plate where testing was terminated was recorded. The total number of errors committed was subtracted from this, and the remainder was the child's raw score.

Group Test Situation: The plates were shown to each of the four classes as a group. They were transferred to 35 mm transparencies and projected onto a 60" x 60" screen. One examiner operated the projector; one called out the appropriate words--each word was given twice; and the remaining four examiners served as proctors to insure that the children were starting at the proper place and that each understood the directions.

Group testing was conducted in the following manner. The projector and screen were prepared and an examiner drew a facsimile of the answer sheet on the chalk board. He than said:

> "I am going to show you some pictures. They will appear here on this screen. There will be four pictures on each slide; one, two, three, and four (pointed to the four appropriate positions [illegible] you see that there are three blocks like this one (pointing to illustration on chalk board) on the half sheet of paper attached to your answer sheets. Mrs. Brooks will say a word; I want you to look at the picture on the screen and place an X in the block which shows the picture that best tells the meaning of the words she says. Let's try some."

Examples A, B, and C were then presented, and the proctors helped each child who indicated lack of understanding of or confusion in following the directions. When it was ascertained that all children in the class understood the procedure, the examiner said:

> "Good, now I'm going to show you some more pictures, and Mrs. Brooks will say the words. I want you to look carefully at all four pictures and place an X over the position of the picture that best tells the meaning of the word Mrs. Brooks says. As we get further along, you will find that you aren't sure you know the word. But I want you to look carefully at all the pictures anyway and make the best choice you can. Remember to look at all the pictures."

Testing was then begun with plate No. 45 and proceeded through plate No. 120. Plates No. 45 through 69 were presented for a period of 10 seconds; plates No. 70 through 120 for 15 seconds. From plates No. 80 through 120 the examiner admonished with each fifth plate, "Be sure to look at all the pictures," or "Be sure to look at all four pictures and make the best choice you can."

Scoring of group answer sheets was begun at plate No. 70 and a basal established in the same manner as for individual testing; that is, scoring progressed upward from No. 70 to point of first error. If eight successful identifications were made before an error was committed, scoring continued until six errors were committed out of eight consecutive presentations. If an error was committed before eight successful identifications had been made, scoring proceeded backward from point of first error to point where eight consecutive successful identifications were obtained, and this point was considered basal. No errors were counted below this point. The terminal point was where six errors were committed out of eight consecutive presentations. No credit was given for successful identification beyond this point. The raw score for each child was computed in the same manner as the raw scores for individual presentations.

No child completed the group tests without a ceiling, or terminal point, being established before plate No. 120 was reached. It was necessary to establish basal points for six children by individual presentation. This was done on the following school day. The greatest number of individual plates to be presented to any of these six children was four.

Testing Times: About 10 minutes are required for individual presentation, and approximately 30 minutes for group testing.

Analysis and Discussion

Means and standard deviations for all sixty subjects on each form, under each type of administration, and for each order were determined. These data are summarized in Table 2.

TABLE 2

PPVT PERFORMANCE OF SIXTY FIFTH GRADE CHILDREN UNDER EXPERIMENTAL CONDITIONS

	Form		Administration		Order	
	A	B	Individual	Group	1st	2nd
Mean	76.57	76.75	77.08	76.23	75.97	77.35
Standard Deviation	9.90	10.32	10.39	9.81	9.64	10.52

From Table 3 it can be seen that neither form of the test, nor type of administration, nor testing order, or any combination of these factors produced a change in PPVT scores which could not be attributed to chance at the .05 level of significance. The obtained F ratio for the form A vs form B comparison was so small as to leave little doubt as to the equivalence of the two forms for the population represented by the group tested. Similarly, an F ratio greater than that obtained in the individual vs group administration comparison would be obtained more than 50 per cent of the time by random sampling alone; considerable confidence in the equivalance of scores obtained by individual and group methods seems warranted.

While an F ratio greater than that obtained between first and second administrations of the PPVT could be expected between 10 and 20 per cent of the time by chance alone, the obtained difference of 1.38 raw score points was in the direction of a practice effect. Consequently, it might be concluded that the practice effect, if any, amounted to only about one point.

TABLE 3

ANALYSIS OF VARIANCE

Source of Variation	d.f.	x^2	s^2	F	$F_{.95}$	$F_{.90}$	$F_{.80}$
Within Subjects	60	1,490.50					
Form (F)	1	1.01	1.01	.04	4.00	2.80	1.68
Administration (A)	1	21.67	21.67	.86	4.00	2.80	1.68
Order (O)	1	57.41	57.41	2.29	4.00	2.80	1.68
F x A x O	1	5.21	5.21	.21	4.00	2.80	1.68
Error (w)	56	1,405.20	25.09				
Between Subjects	59	10,576.49					
F x A	1	429.41	429.41	2.44	4.00	2.80	1.68
F x O	1	143.00	143.00	.81	4.00	2.80	1.68
O x A	1	147.41	147.41	.84	4.00	2.80	1.68
Error (b)	56	9,856.67	176.01				
Total	119	12,066.99					

Conclusions

When both forms of the PPVT were administered in counterbalanced order to fifth grade children of normal intelligence and in the proper school grade for their age, their average scores were neither a function of the form administered nor of having the test administered individually or in a group. Furthermore, any practice effect which might have stemmed from having had the alternate form two days earlier averaged only about one point.

Reference

1. Dunn, L. M., Peabody Picture Vocabulary Test Manual. Tentative edition. Nashville, Tennessee: George Peabody College for Teachers, 1958.

2. Lindquist, E. F., Design and Analysis of Experiments in Psychology and Education. New York: Houghton Mifflin Company, 1953.

APPENDIX D

PEABODY PICTURE VOCABULARY TEST PERFORMANCE OF EDUCABLE MENTALLY RETARDED CHILDREN[1]

Lloyd M. Dunn and Sadye Brooks
George Peabody College for Teachers

The present study was conducted to provide preliminary data on the utility of the Peabody Picture Vocabulary Test (PPVT) as a psychometric instrument for measuring the verbal intelligence of educable mentally retarded children and youth. This untimed test of auditory comprehension for the spoken word is individually administered. It consists of 150 plates containing four pictures each. This one set of plates provide the test materials for both Form A and B of the scale. The plates are arranged in ascending order of difficulty, empirically-determined from data gathered in the standardization of the instrument. To administer the PPVT, the examiner provides a stimulus word orally, such as "Show me arrow". The subject then points to, or otherwise indicates, the picture on the plate which best illustrates the meaning of the word. Each subject is tested from a basal of eight consecutive correct responses to a ceiling of six errors in eight consecutive responses. The raw score for the subject is his basal score plus the number of correct responses up to his ceiling. A more complete description of the instrument is reported by Dunn (1) in his manual.

While the scale was standardized on random samples of subjects of all levels of intelligence from 2 years 6 months through 18 years, it was designed to be especially effective in studying the mentally retarded. Since the line drawings are fairly large and relatively free of figure ground perception problems, it is made more applicable to brain injured children. Since a pointing response is required rather than an oral one, it is especially appropriate for subjects who have speech difficulties or who are extremely withdrawn. Since no reading is required, it has advantages in testing children who are inferior or non-readers. Since the test is untimed, it is especially suited for slow subjects and those who are multiply handicapped.

Before one can assume the PPVT is an adequate psychometric measure for the mentally retarded, evidence as to its reliability and validity with such subjects must be demonstrated. To obtain this information the following analyses were made:

---

[1] The research reported herein was performed pursuant to Contract SAE-6425, United States Office of Education, Department of Health, Education, and Welfare.

Appreciation is extended the Nashville and Davidson County School Systems.

1. A comparison was made of scores obtained by retarded subjects on Forms A and B of the PPVT. (Means and standard deviations of raw scores, vocabulary age scores, and vocabulary quotient scores were calculated and compared. In addition the correlation between Form A and B vocabulary age scores was derived to measure reliability.)

2. The relationship was studied between order of difficulty of the test items for the retarded subjects and the order of presentation (which constituted the order of difficulty for the individuals on whom the test was standardized). A comparison was also made of differences in item difficulty between Forms A and B.

3. Age equivalents were compared on the PPVT and the 1937 Revised Stanford-Binet Tests of Intelligence, Form L. In addition the correlation between VA's on the PPVT and MA's on the RS-B was derived as evidence for the validity of the PPVT.

4. Vocabulary quotients on the PPVT were compared with intelligence quotients on the RS-B.

5. The relationship was studied between vocabulary age scores on the PPVT and chronological ages of the retarded subjects.

Procedure

One hundred thirty white, educable mentally retarded pupils attending the city and county special day classes in the environs of Nashville, Tennessee constituted the sample for this study. Of this number 84 were males and 46 females. Each pupil was given the Revised Stanford-Binet, Form L, prior to his admission to the special class and scores were recorded in cumulative folders. These test results were used in the present study, the RS-B having been administered to the pupils one to three years prior to the PPVT. Table 1 shows the mean and standard deviation of CA's, MA's, and IQ's for the group. The subjects ranged in chronological age from 7-0 to 15-11, with a mean of 12-10 and a standard deviation of 2-3. The IQ's ranged from 50 to 80 with a mean of 66 and a standard deviation of 7.03. Mental ages were computed as of the day of administration of the first form of the PPVT. These MA's ranged from 3-10 to 11-2 with a mean of 8-4 and a standard deviation of 1-7.

TABLE 1

BACKGROUND DATA ON THE SUBJECTS

	Chronological Ages	Stanford-Binet Mental Ages	Stanford Binet Intelligence Quotients
Mean	12-10	8-4	66
Standard deviation	2-3	1-7	7.03

Both Forms A and B of the PPVT were administered individually to all 130 subjects. These forms were administered in counterbalanced order to offset the accumulation of bias due to practice.

Results and Discussion

Table 2 shows the means and standard deviations of PPVT scores. Besides raw scores, data on vocabulary age and vocabulary quotient are summarized. (Vocabulary age norms are comparable to mental age scores on the Binet, and were derived in a similar manner. Vocabulary quotient scores are somewhat comparable to Stanford-Binet IQ scores but were calculated as standard score equivalents in the PPVT standardization giving a separate set of derived scores at each age level and using a mean of 100 and a standard deviation of 15.)

TABLE 2

SUMMARY TABLE OF PEABODY PICTURE VOCABULARY TEST SCORES

	Raw Scores		Vocabulary Ages		Vocabulary Quotients	
	Form		Form		Form	
Measure	A	B	A	B	A	B
Mean	65.51	65.07	8-0	7-10	69.9	69.4
Standard deviation	10.36	9.28	2-0	1-8	10.2	8.6

Table 3 gives correlations between test score results.

TABLE 3

INTERCORRELATION OF TEST RESULTS

Comparisons	Correlations*
A. Reliability Data	
VA (Form A) vs VA (Form B)	0.76
Order of difficulty vs Order of administration (Form A)	0.99
Order of difficulty vs Order of administration (Form B)	0.99
Order of difficulty (Form A) vs Order of difficulty (Form B)	0.98
B. Validity Data	
VA (Form A) vs MA (RS-B)	0.74
VA (Form A) vs CA	0.68

* All correlations significant at the 0.01 level.

A. Reliability Data

1. Differences between means were minute and between standard deviations were slightly larger on Forms A and B of the PPVT. Mean raw scores on Form A and B differed by less than one point; the mean Form A VA was two months higher than that for Form B; the mean VQ-IQ difference was only one-half a point. The difference between standard deviations was 4 months for VA's and 1.6 points for VQ-IQ's. The product-moment correlation between VA scores for Forms A and B was 0.76, indicating a reasonably reliable test for educable mentally retarded pupils of school age.

2. In standardizing the PPVT, the 150 plates (and thereby the items) were arranged in ascending order of difficulty, empirically determined from test results obtained on a large sample of subjects of all levels of intelligence. This same order of difficulty appears to hold for the educable mentally retarded. Using the rank order correlation technique, a comparison was made between the order of difficulty of items found for the educable group and the order of administration (the order of difficulty for persons across the entire range of intelligence) yielding a correlation of 0.99 for both Forms A and B. The correlation and order of item difficulty between Form A and B was 0.98. These extremely high positive correlations indicate the items on both Forms A and B of the PPVT are in order of difficulty not only for intellectually normal subjects but also for educable mentally retarded individuals of school age. This again adds support to the contention that the PPVT is a reliable instrument for use with the educable group.

3. The difference between PPVT and RS-B mean age equivalents was four months for Form A and six months for Form B, the RS-B MA's being higher. Standard deviations of age equivalent scores was higher for VA's on the PPVT than for MA's on the RS-B, being five months for Form A and one month for Form B.

4. Vocabulary quotients on the PPVT were, on the average, slightly above intelligence quotients on the RS-B. The mean difference was 3.9 points for Form A and 3.4 points for Form B. This tendency for VQ's to be above IQ's while VA's were below MA's is possible since they are different types of derived scores. IQ's were derived from MA's using the IQ = MA/CA formula; VQ's were derived as standard score equivalents using a mean of 100 and a standard deviation of 15. The standard deviations of VQ scores was slightly larger than that for IQ scores, the difference being 3.17 for Form A and 1.57 for Form B. A correlation coefficient was not found since it would have been of the same order as that for age equivalents.

5. There was a fairly close relationship between vocabulary age scores on the PPVT and chronological ages. The product moment correlation for VA (Form A) vs CA was 0.68. The increases in vocabulary ages with successive increases in chronological ages is indirect but not conclusive evidence for the validity of the PPVT. Physical strength, ossification, height, and weight also increase with age but are not highly related to verbal intelligence.

Summary and Conclusions

Forms A and B of the Peabody Picture Vocabulary Test were administered in counterbalanced order, to 130 white, educable mentally retarded pupils in the city and county public day schools of Nashville, Tennessee. Revised Stanford-Binet, Form L, scores were also available, this instrument having been administered when the pupils were placed in special classes. An analysis of these test scores showed a close correspondence between alternate forms of the PPVT. It was also found that PPVT scores compared fairly closely with the Stanford-Binet scores, PPVT vocabulary ages being, on the average, five months lower than RS-B.

Until evidence is accumulated to the contrary, it would appear that the PPVT may be used effectively to measure the verbal intelligence of educable mentally retarded pupils of school age. The reliability coefficient was found to be 0.76 and the validity coefficient 0.74. The test is quickly administered to the educable mentally retarded in 5 to 10 minutes, with no modifications in test procedure. The test was found to be especially effective with multiple handicapped subjects. Further investigation of the utility of the PPVT is desirable to add to the preliminary data reported in this article.

References

1. Dunn, Lloyd M., Peabody Picture Vocabulary Test Manual. Tentative edition. Nashville, Tennessee: George Peabody College for Teachers, 1958.

APPENDIX E

PEABODY PICTURE VOCABULARY TEST PERFORMANCE OF TRAINABLE MENTALLY RETARDED CHILDREN[1]

Lloyd M. Dunn and John V. Hottel
George Peabody College for Teachers

Under contracts with the Tennessee State Department of Education and the U. S. Office of Education, Peabody College conducted a study in 1955-57 on "the effectiveness of special day class training programs for severely (trainable) mentally retarded children (1)". Early in the investigation it became apparent that most trainable children were especially handicapped in the area of speech development. Since many of these youngsters lacked intelligible speech, they were penalized on such tests of verbal intelligence as the Revised Stanford-Binet. Yet the investigators felt that many of these children understood the meaning of many more words than they were able to produce. Because none of the existing vocabulary scales was found to be sufficiently sensitive to measure relatively slight increments of growth in auditory comprehension for the spoken word, a project was inaugurated at Peabody College to develop a vocabulary test which would be especially useful for studying children with IQ's of 30 to 50.

The resulting test was the Peabody Picture Vocabulary Test (PPVT). It consists of 150 plates, each composed of four pictures. This one set of plates provides the test material for both Form A and Form B of the scale. To administer the PPVT, the examiner provides orally a stimulus word i.e. "Show me arrow". Then the subject points to, or otherwise indicates, the picture on the plate which best illustrates the meaning of the word i.e. arrow. Each subject is tested from a basal of eight correct responses to a ceiling of six errors in eight consecutive responses. The raw score for the subject is his basal score plus the number of correct responses up to his ceiling.

To make the test of maximum utility, it was standardized on a random sample of subjects of all levels of intelligence over the age range 2-6 through 18-0 years. The two major reasons for this action were: (1) to provide a better frame of reference for studying the trainable child since they could be fitted into the total spectrum of intelligence, and (2) to provide derived scores comparable to those obtained from such tests as the Revised Stanford-Binet and the Vineland

---

[1] The research reported herein was performed pursuant to Contract SAE-6425, United States Office of Education, Department of Health, Education, and Welfare.

The assistance of the teachers of trainable children throughout Tennessee is gratefully acknowledged.

Social Maturity Scale (age equivalents (vocabulary ages) and standard score equivalents (vocabulary quotients) were derived in standardizing the instrument).

A more complete description of the instrument and its standardization is reported by Dunn (2) in his manual.

Purpose

It was the purpose of the present study to investigate the usefulness of the Peabody Picture Vocabulary Test as a psychometric device for measuring the verbal intelligence of "trainable" mentally retarded children through gathering data on its reliability and validity. More specifically this investigation was designed:

A. Reliability Data

1. To provide reliability data through comparing scores on Forms A and B of the test;

2. to compare observed order of difficulty of the test items for trainable children with the order in which they were presented to the children (which constituted the order of difficulty of the items for the individuals on whom the test was standardized since they were placed in ascending order of difficulty at that time);

B. Validity Data

3. to study the relationship between vocabulary age and vocabulary quotient scores on the PPVT and mental age and intelligence quotient scores on the RS-B;

4. to compare scores obtained by trainable children on the PPVT versus teacher ratings of their ability to read;

5. to compare scores obtained by trainable children on the PPVT versus teacher ratings of their ability to produce written language; and

6. to study the relationship between vocabulary age scores on the PPVT and chronological ages.

Procedure

Forms A and B of the Peabody Picture Vocabulary Test were administered, in counterbalanced design, to 364 "trainable" children enrolled in 29 special day classes operated by the public schools throughout the State of Tennessee. The investigation was restricted to include only children with chronological ages between 6-0 and 15-11. Background data on the subjects are summarized in Table 1.

TABLE 1

BACKGROUND DATA ON THE SUBJECTS

	Chronological Ages	Stanford-Binet Mental Ages	Stanford-Binet Intelligence Quotients
Mean	11-4	4-6	44.8
Standard deviation	2-4	0-5	8.2

The mean CA of the group was 11-4 with a standard deviation of 2-4. Intelligence quotients of the children were obtained from the 1937 Revised Stanford-Binet Tests of Intelligence, Form L, which was administered to each child when he was enrolled in the special class one to two years prior to the PPVT. On the assumption that IQ remains constant, mental ages (MA's) were computed on the basis of CA at the time of the administration of the PPVT. The IQ's of the children ranged from 19 to 59 with a mean of 44.8 and a standard deviation of 8.2. Mental ages ranged from 1-8 to 8-3 with a mean of 4-6 and a standard deviation of 5 months.

The etiologies of defect or clinical types of mental deficiency represented by members of the group were classified into four general categories. The category labeled "Brain Injury" included all children whose defect had been diagnosed as a result of organic injury to brain tissue. It included children with and without cerebral palsy; those who were under sedation for the control of seizures and those who had never had sedation and/or who displayed no signs of convulsive attack; and both hyperactive and hypoactive children. In short, this was a most inclusive category. The classification of "Cranial Anomalies" included hydrocephalics, microcephalics, and one oxycephalic. The classification of "Thyroid Dysfunction" included one hypothyroid child (a cretin) and one hyperthyroid child. This distribution along with the percentage of each classification is presented in Table 2.

TABLE 2

DISTRIBUTION BY ETIOLOGY OR CLINICAL TYPES

Classification	Number	Per cent of Total
Total	364	100
Brain injury	165	45
Mongoloid	125	34
Unknown or unclassified	66	18
Cranial anomalies	6	2
Thyroid dysfunction	2	1

The 29 teachers of "trainable" children throughout the State of Tennessee were taught how to administer and score the PPVT. This was accomplished by first calling all the teachers together for a two-day workshop in Nashville, Tennessee. Following this, the investigators visited each of the classes and checked out each of the teachers by again demonstrating the administration of the test, and then observing them while they administered the instrument. The teachers were then left to administer both forms of the test to each of their pupils in a two-week period. Since the teachers had attendants, they were able to take each pupil in turn into an ante-room to administer the scale.

In addition to administering the PPVT, each teacher was asked to rate each of his children on a point-scale estimate of both reading and written language achievement. Not all teachers rated all of their pupils. The categories and the number of pupils rated at each are summarized in Table 3 below.

TABLE 3

TEACHER RATINGS OF LANGUAGE ARTS ACHIEVEMENT

Number of pupils rated	Category
	A. Reading Achievement
56	Does not yet recognize his first name (when placed on a flashcard and mixed with those of his classmates).
84	Recognizes his first name (when placed on a flashcard and mixed with those of his classmates).
13	Recognizes his surname, in isolation, in print.
93	Recognizes the first names of 3 or more of his classmates (when placed on flashcards).
35	Has a sight reading vocabulary of 5-15 words other than classmates' names (when placed on flashcards).
6	Has a sight reading vocabulary of 16-30 words.
6	Has a sight reading vocabulary of 31-50 words.
5	Reads at the beginning grade 2 level.
3	Reads at the beginning grade 3 level.
301	

TABLE 3 (continued)

TEACHER RATINGS OF LANGUAGE ARTS ACHIEVEMENT

Number of pupils rated	Category
	B. Written Language Achievement
92	Does not yet copy his own first name.
42	Copies his own first name.
49	Prints (or writes) his own first name from memory.
38	Prints (or writes) his first and surname from memory.
19	Prints (or writes) 3-5 words other than names of classmates from memory.
7	Prints (or writes) simple sentences of five or more words from memory
247	

Relationhsips between different sets of scores were measured by either the Pearson product-moment or the Spearman rank-order correlational techniques. Resulting coefficients are summarized in Table 4.

Results and Discussion

A. Reliability Data

1. The product-moment correlation between VA scores for Form A and B was 0.87. This highly significant coefficient indicates that the PPVT is a reasonably reliable test for trainable mentally retarded pupils of school age. It was assumed with this high degree of reliability, that relationships between Form A score and other variables would be essentially the same as those between Form B scores and the same variables. Therefore, only comparisons with Form A scores were computed.

2. In standardizing the PPVT, the 150 plates (and thereby the items) were arranged in ascending order of difficulty, emperically determined from test results obtained on a large sample of subjects of all levels of intelligence. This same order of difficulty appears to hold for trainable children. Using the rank order technique, a correlation of 0.98 was found between order of difficulty and order of presentation for Form A, and 0.99 for Form B. These findings provide further evidence that the PPVT is a reliable instrument for trainable children.

TABLE 4

INTERCORRELATION OF TEST RESULTS

Comparison	Number	Correlation*
A. Reliability Data		
1. VA (Form A vs VA (Form B)	364	0.87
2. Order of difficulty vs order of administration (Form A)	364	0.98
Order of difficulty vs order of administration (Form B)	364	0.99
B. Validity Data		
3. VA (Form A vs MA (RS-B)	267	0.71
4. VA(Form A vs reading achievement	309	0.44
5. VA (Form A) vs written language achievement	261	0.50
6. VA (Form A) vs CA	364	0.47

* All correlations significant at the 0.01 level.

B. Validity Data

3. The product-moment correlation between PPVT, Form A, vocabulary age scores and RS-B, Form L, mental age scores was 0.71. This significant positive coefficient indicates satisfactory validity for the PPVT when measured against the RS-B criteria. The validity of PPVT, VQ's versus RS-B, IQ's could not be calculated for trainable children. The VQ's in the PPVT are standard score equivalents with a mean of 100 and a standard deviation of 15, and were only found for scores out to three standard deviations. Therefore the lowest VQ in the PPVT tables found in the manual is 55. Users of the scale may wish to experiment with VQ's for trainable children calculated from VQ= VA/MA, but it remains for future research to provide evidence as to the utility of this procedure.

4. The PPVT was designed as a measure of verbal intelligence and therefore should be especially suited to predicting school success in the language arts area. While it should be easier to test this contention among "educable" children, it was decided to investigate this relationship among the trainable group. The teacher ratings for reading achievement in Table 4 indicate that over 80 per cent of the trainable children recognize only their own name and those of their classmates in print. It should not be surprising that 3 children were reading at the beginning Grade 3 levels since mental ages of the subjects ran as high as 8-3, IQ's ran up to 59, and CA's to 15-11. Nevertheless, throughout Tennessee, even though some of the children had been in classes for a number of years, first under the sponsorship of parent groups and later

under the public schools, not one trainable child was found to be reading at the point where he could be called literate, if one takes the upper Grade 4 reading level as a critereon.

The product-moment correlation between vocabulary age and reading achievement was 0.44. This positive relationship was found to be statistically significant, but was far from a perfect or even a high one.

5. One would predict that the relationship between vocabulary age scores and written language achievement would be of the same order as that for reading. The product-moment correlation was slightly higher, being 0.50. Here again, about 90 per cent of trainable children could print nothing more complex from memory than their own name. The fact that 7 of the 247 trainable children studied could print simple sentences of five or more words from memory is not unrealistic when one again recalls that MA's ran as high as 8-3, IQ's to 59, and CA's to 15-11. No trainable children were found to be academically educable in the present study.

6. A fair relationship was found between vocabulary age and chronological age. The product-moment correlation was 0.47, a statistically significant one. The increases in vocabulary age with successive increases in chronological age lends only indirect evidence for the validity of the PPVT; physical strength, height, and weight also increase with age but are not highly related to verbal intelligence.

In interpreting the data in this study, it must be kept in mind that trainable mentally retarded children are a fringe group, generally scoring over three standard deviations below the mean on the tests. This factor is of considerable importance in interpreting correlation ratios since non-normal distributions and extreme scores are the rule rather than the exception.

Summary

Forms A and B of the Peabody Picture Vocabulary Test were administered, in counterbalanced order, to 364 trainable mentally retarded children enrolled in the special day classes of the public schools of the State of Tennessee. Revised Stanford-Binet, Form L, scores were also available, this instrument having been administered to the group in the preceding two years. At the time the PPVT scores were obtained, the teachers rated the reading and writing achievement of their pupils on point scales. Correlations between sets of scores were computed.

The correlation coefficient between Forms A and B of the PPVT proved to be 0.87. Strong evidence was found that the difficulty level of the individual items on the scale is the same for trainable children as for persons of all levels of intelligence on whom the test was standardized. These data indicate the PPVT is a very reliable instrument for "trainable" pupils of school age.

Validity data were positive but not as conclusive as that for reliability. The correlation of PPVT vocabulary age scores and RS-B mental age scores was 0.71. The relationship of vocabulary age with reading achievement was 0.44, with written language achievement was

0.50, and with chronological age was 0.47. Nevertheless, until evidence is accumulated to the contrary, the PPVT should be considered reasonably valid as an instrument for measuring the verbal intelligence of "trainable" boys and girls.

Descriptive data gathered in connection with the reliability and validity information indicate that not one trainable pupil studied had become literate (or academically educable) if one accepts the Grade 4 achievement as the point of literacy. Over 80 per cent of the children could recognize in print, at best, their own names, and those of some of their classmates. About 90 per cent could print nothing more complex from memory than their own name.. However three pupils could read at the beginning Grade 3 level, and seven could print simple sentences from memory. This is not unrealistic since MA's ran as high as 8-3, IQ's to 59, and CA's to 15-11.

References

1. Dunn, Lloyd M. and John V. Hottel., The Effectiveness of Special Day Class Training Programs for Severely (Trainable) Mentally Retarded Children. Nashville, Tennessee: George Peabody College for Teachers, 1958.

2. Dunn, Lloyd M., Peabody Picture Vocabulary Test Manual. Tentative edition. Nashville, Tennessee: George Peabody College for Teachers, 1958.

APPENDIX F

COMPARABILITY OF PEABODY, AMMONS, VAN ALSTYNE AND COLUMBIA TEST SCORES WITH CEREBRAL PALSIED CHILDREN[1]

Lloyd M. Dunn and Randall K. Harley
George Peabody College for Teachers

Appraisal of the intelligence of cerebral palsied children has passed through three stages. Early efforts involved clinical judgements based on non-standardized behavioral observations. Then followed the stage when modifications of Binet and Wechsler scales were widely employed wherein test procedures were varied from subject to subject; of course, the use of norms was, at best, hazardous. Even more recently, new-type tests have begun to appear which are being administered in a similar manner to subjects with and without speech and motor disabilities. These instruments have facilitated the collation of test results on groups of cerebral palsied subjects since standard procedures could be used for the first time.

Purpose

The purpose of the present study was to investigate the comparability of four new-type individual tests of intelligence with cerebral palsied children. One test -- the Peabody Picture Vocabulary Test (PPVT) -- is completely new to the field. Two others -- the Van Alstyne Picture Vocabulary Test and the Columbia Mental Maturity Scale -- have been extensively revised. The Ammons Full-Range Picture Vocabulary Test is the only instrument used in this study which has been available for some time in its present form.

More specifically this investigation was designed to explore:

1. The distribution of mental age scores obtained by cerebral palsied subjects on the four instruments;

2. intercorrelations among the four tests as well as correlations between alternate forms of the PPVT and Ammons;

3. correlations of the four tests with teacher ratings of reading and arithmetic achievement.

---

[1] The research reported herein was performed pursuant to Contract SAE-6425, United States Office of Education, Department of Health, Education, and Welfare.

Special appreciation is acknowledged Harold F. Bligh of the World Book Company for making the Van Alstyne and Columbia scales available to the investigators prior to publication date. The assistance of the Peabody College doctoral students in psychology is also recognized.

Instruments

The Peabody (3), Ammons (1), and Van Alystyne (4) Picture Vocabulary Tests are somewhat similar in design. All three consist of a series of rectangular plates or cards, containing four pictures each. All are individual, power tests of hearing vocabulary and/or verbal intelligence -- using oral instructions, pictorial materials, and the pointing (or "yes" or "no") response. To administer the scales, the examiner provides a stimulus word orally as in "Show me ball". The subject then points to, or in some other way, indicates the picture on the plate which best illustrates the meaning of the stimulus word. With severely-involved, cerebral palsied subjects who are unable to point or evoke an oral "yes" "no" response, the examiner may point to the four pictures in order on a card, obtaining a "yes" or "no" response for each item in turn.

The Peabody Picture Vocabulary Test (PPVT) is a new test developed by Dunn. It consists of 150 plates arranged in ascending order of empirically determined difficulty extending over the age range 2-6 through 18-0, but having extrapolated age norms below this level. The one set of plates serves for both Forms A and B of the test. For each form, one stimulus word only is illustrated by a plate. The subject is tested from a basal of eight consecutive plates. The raw score for a subject is his basal score plus the number of items answered correctly up to his ceiling. Age, standard score, and percentile norms are provided, based on a standardizing population of 4,012 children and youth.

The Revised Van Alystyne Picture Vocabulary Test was developed by the World Book Company in cooperation with Van Alystyne. In revising this well-known pre-school test which was first published in 1929, completely new pictorial materials are provided. Thirty-two of the words used in the original edition have been retained. Sufficient new words or phrases have been added to bring the stimulus words up to a total of 60. Since one plate is used to test only one item, the test consists of 60 plates. The test is primarily designed for children with mental ages between two and five years, though the age norms have been extrapolated from 2-1 to 10-5. There is only one form of the test. The raw score is obtained by counting the number of correct responses obtained by the subject when the examiner sweeps through the odd numbered of the 60 plates and then the even numbered.

Since its publication in 1948 by Ammons, The Full-Range Picture Vocabulary Test has been recognized as an especially useful instrument for testing cerebral palsied children. The unique feature of this instrument is the efficient use made of only 16 cards, each consisting of the usual four line drawings. Each card is used to illustrate a series of words of increasing difficulty, rather than just one word per form as on the other tests. Approximately 85 test words are illustrated in each of the two parallel forms by the single set of 16 cards. Another difference between the Ammons and other tests is in relation to guessing. The authors do not allow the subject to guess, asking him to define words, or point again later in the test to words the examiner thought were guessed by the subject. The test measures over the mental age range 2-6 through adulthood. As in the case of the Van Alystyne only age equivalents are provided for children.

The Revised Columbia Mental Maturity Scale (2) was developed by Burgemeister, Blum, and Lorge in cooperation with the World Book Company. This instrument differs from the picture vocabulary tests in that it does not measure auditory comprehension for the spoken word but rather the ability to select the drawing which does not belong with the others. The items in the lower half of the scale are based on concepts such as color or simple geometric form, but at the upper end, recognition of very subtle relations in pairs of pictures is necessary to classify the pictures so the one which does not belong can be found. The test consists of 100 elongated plates, each containing 3 to 5 drawings in a row. In revising it, 17 of the 100 plates and in the original edition have been replaced, and the order of many of the remaining plates has been re-arranged so as to place the plates in ascending order of difficulty, based on new evidence accumulated in re-standardizing the instrument. The test is designed for children from 3 to 12 years of age. Mental age norms extend over the range 3-5 to 13-11. There is only one form of the test. Raw scores are found by scoring the correct responses in sweeping upward through the 100 plates until the child fails twelve in any sequence of sixteen cards.

Procedure

The subjects used in the study were all 20 of the cerebral palsied children attending a three-class special unit for crippled children in one of the Nashville City Schools. There were 10 boys and 10 girls in the group. The children were classified as 10 athetoid, 9 spastic, and 1 ataxia. Degrees of involvement were 7 severe, 6 moderate, and 7 mild. Three of the group had appreciable hearing losses and two had moderate vision defects. Three had no speech, 5 poor, 5 fair, and 6 average speech for age. Eight were non-ambulatory, 6 walked with support, and 6 walked without support. Fifteen of the subjects used a pointing response while 5 indicated "yes" or "no" as the examiner pointed out the alternates in order. The mean chronological age of the group was 11-0 with a standard deviation of 34.35 months, and a range of 7-1 through 16-2.

The Peabody, Ammons, Van Alystyne and Columbia Tests were administered in a period of two weeks to the children. The PPVT and Ammons were given during the first week by fourteen advanced graduate students in psychology -- with tests, forms of tests, subjects, and examiners counterbalanced. The original design called for the Van Alystyne and Columbia to be administered at the same time as the PPVT and Ammons, with all four tests counterbalanced for order of presentation. However, the Van Alystyne and Columbia did not arrive in time and had to be administered in the second week by two examiners in counterbalanced order. Subjects were not given two scales at one sitting. For all 20 subjects, the test was administered and scored according to procedures prescribed in the manuals. Raw scores were then converted to age equivalent norms.

In addition to obtaining test scores, the teachers were asked to rank the 20 cerebral palsied children on reading and arithmetic achievement. Since all three teachers knew all the children it was possible for them to reach concensus in rank ordering. (In arithmetic the teachers estimated a range from a high of Grade 5 achievement to a low of no comprehension of quantitative concepts; the range of reading achievement extended from Grade 7 to the low reading readiness level.)

Table 1 summarizes the test scores and chronological ages of all 20 cerebral palsied children.

TABLE 1

CHRONOLOGICAL AGES AND TEST SCORES OF THE 20 CEREBRAL PALSIED CHILDREN

			Mental Ages						Rank	
			PPVT		Ammons					
Child	CA	Md. M.A.	A	B	A	B	Van Alystyne	Columbia	Rdg. Ach.	Arith. Ach.
1	16-2	11-11	11-4	13-4	12-6	13-6	8-0	7-1	1	3
2	12-8	10-10	11-0	10-10	13-3	13-9	9-7	10-0	6	2
3	8-10	10-6	10-7	10-5	11-6	10-9	9-8	7-4	9	8
4	13-8	8-11	10-5	8-4	9-6	11-0	8-4	7-0	5	4
5	13-8	8-10.5	8-9	9-1	9-8	11-3	7-6	6-6	4	5
6	11-5	8-10	8-7	7-5	10-9	9-4	7-0	9-1	2	6
7	11-9	8-1.5	8-3	8-0	10-4	10-3	7-0	6-2	11	10
8	11-0	7-9.5	6-10	7-3	9-6	8-6	6-1	8-4	7	7
9	15-3	7-7	7-3	7-5	7-3	10-0	7-9	9-8	3	1
10	8-11	7-2	6-10	7-0	8-0	10-6	7-3	7-1	8	9
11	11-0	6-8	6-1	7-3	7-9	8-2	5-1	5-8	13	13
12	7-6	6-4.5	7-3	6-6	6-3	7-0	5-9	5-3	12	12
13	11-11	6-2.5	4-10	6-6	7-3	8-2	5-9	5-11	10	11
14	11-1	6-1	5-7	5-1	8-9	3-8	4-8	5-1	14	14
15	14-8	4-11	4-5	3-10	7-3	7-0	4-11	4-11	15	15
16	7-3	4-6	4-3	4-2	4-9	7-6	5-1	3-11	16	16
17	7-5	3-11	3-9	3-11	4-1	3-11	4-0	3-6	17	17
18	7-1	3-2	2-7	2-6	3-8	2-10	3-6	3-9	18	18
19	7-1	2-8	1-6	1-10	2-6	2-10	3-5	4-2	19	19
20	11-3	2-7	2-3	2-1	2-6	3-2	2-8	3-9	20	20
Mean	11-0		6-8	6-8	7-10	8-2	6-2	6-3		
Standard deviation	34.35		35.73	36.35	37.99	41.00	23.71	23.95		

Results and Discussion

1. The mean MA scores, as seen in Table 1, ranged from a low of 6-2 on the Van Alystyne to a high of 8-2 on the Ammons. From Table 2 it can be seen that a significant F ratio was found indicating that

differences in mean MA scores could not be attributable to chance at the 0.05 level of confidence. A Hartley test was run giving an F ma of 2.99 which did not fall in the critical region. When t-tests were then used on differences between pairs of means, significant results occurred, between three sets of scores, as shown in Table 3.

TABLE 2

ANALYSIS OF VARIANCE FOR MEAN MENTAL AGE SCORES

Source of variation	Sum of squares	d.f.	Mean square	F	$F_{.95}$
Individuals	112,442.76	19			
Tests	10,629.54	5			
Error	17,844.29	95	2,125.91	11.32*	2.72
Total	140,916.59	119	187.83		

* significant at 0.05 level of confidence.

TABLE 3

SUMMARY OF T-TESTS TO DETERMINE SIGNIFICANCE BETWEEN PAIRS OF MENTAL AGE SCORES

	PPVT		Ammons		Van Alystyne	Col-umbia
	A	B	A	B		
PPVT						
Form A		0.02	1.23	1.48	0.66	0.56
Form B			1.24	1.49	0.63	0.53
Ammons						
Form A				0.29	2.06*	1.96
Form B					2.29*	2.20*
Van Alystyne						0.13

* significant at the 0.05 level of confidence

The students scored significantly higher on Form A of the Ammons than on the Van Alystyne; their scores on Form B of the Ammons were significantly above both their scores on the Van Alystyne and Columbia (see Table 1).

Examination of the descriptive data in Table 1 indicates the following: (1) MA's on the PPVT from 2-0 to 4-5 were consistently lower than those on the other tests. This may be because the PPVT has more bottom than the other instruments or may an artifact of the standardization population. (2) MA's on the Van Alystyne and the Columbia were consistently lower from 8-0 and up than scores on the other two tests. In the case of the Van Alystyne, this is probably due to lack of top; the test was only intended for children with mental ages between two and five years. The depressed scores on the Columbia cannot be so easily explained since the test is intended for children from 3 to 12 years. Perhaps cerebral palsied children have special difficulty with the abstract pairing and discrimination processes required at the upper levels of this test.

2. Correlations between sets of mental age test scores are summarized in Table 4. These were computed by the Spearman rank-order technique. As was expected, because of the wide range of ages of the 20 subjects, these coefficients were quite high. No doubt, with a more homogenous group of cerebral palsied children, they would not have been so large.

TABLE 4

CORRELATIONS BETWEEN PAIRS OF MENTAL AGE TEST SCORES

	PPVT		Ammons		Van Alystyne	Columbia
	A	B	A	B		
PPVT						
Form A		.97	.91	.92	.94	.82
Form B			.92	.95	.94	.82
Ammons						
Form A				.86	.88	.80
Form B					.96	.81
Van Alystyne						.87

The alternate form reliability coefficient for the PPVT was 0.97 and for the Ammons 0.86. Both tests appear to have sufficient consistency to justify their use with cerebral palsied children.

The intercorrelations between tests generally range around 0.90 for the picture vocabulary scales. However these scales only correlate in the vicinity of 0.80 with the Columbia. This probably indicates that the Columbia may be measuring intellectual factors other than these tapped by the picture vocabulary tests.

3. Intercorrelations of teacher rankings in arithmetic and reading achievement with mental age test scores are summarized for all four tests in Table 5. These were also computed by the rank order technique. All are relatively high positive correlations. It had been speculated that Columbia scores would correlate much higher with arithmetic than with reading achievement ranks; and that vocabulary scale scores would correlate higher with reading than with arithmetic achievement ranks. No such tendency is discernible. The only slight emphasis there might be, is that all tests tended to correlate more closely with arithmetic ranks than with reading ranks.

TABLE 5

INTERCORRELATIONS OF MENTAL AGE TEST SCORES AND TEACHER RANKINGS IN READING AND ARITHMETIC ACHIEVEMENT

Area	PPVT		Ammons		Van Alystyne	Col-umbia
	A	B	A	B		
Reading	.87	.87	.80	.86	.87	.89
Arithmetic	.90	.91	.80	.90	.93	.93

Summary

The present study was undertaken to explore the usefulness with cerebral palsied children of three new or revised individual tests of intelligence and their comparability with a well-established test. The new Peabody, the 1948 Ammons, and the Revised Van Alystyne Picture Vocabulary Tests; and the Revised Columbia Mental Maturity Scale were administered to twenty children with various types and degrees of cerebral palsy. Teacher rankings of reading and arithmetic achievement were also obtained on the boys and girls.

The PPVT would appear to have more bottom than any of the other tests, and the Columbia the least. Older cerebral palsied children obtained lower mental ages on the Van Alystyne and Columbia than on the other tests, probably indicating insufficient ceiling especially for the Van Alystyne. Mean Ammons mental age scores were significantly higher than those on the Van Alystyne and Columbia.

The alternate form reliability coefficient was found to be 0.97 for the PPVT and 0.86 for the Ammons. Intercorrelations between sets of mental age scores on the four tests were all found to exceed 0.80. However picture vocabulary test scores correlated more closely with one another than with the Columbia, indicating, as was expected, that the latter may be measuring different intellectual factors than the vocabulary scales.

Intercorrelations of teacher rankings of reading and arithmetic achievement and mental age test scores were 0.80 and above for all four tests. Contrary to expectancy, there was a tendency for all tests to correlate more closely with arithmetic than with reading. It had been anticipated that only the Columbia would do this.

The results of this study indicate that all four tests can be used successfully with cerebral palsied children in predicting school success. When it is desirable for mental age scores to parallel those for other tests, the range for the Van Alystyne will probably be most effective for children with mental ages 3-0 to 8-0, and the Columbia for children whose mental ages fall between 4-0 and 9-0. Both the Revised Van Alystyne Picture Vocabulary Test and the Revised Columbia Mental Maturity Scale appear to be much more satisfactory than the original versions.

References

1. Ammons, R. B, and Ammons, H.S. The Full-Range Picture Vocabulary Test. New Orleans: R. B. Ammons, 1948. (now available from Psychological Test Specialists, Box 1441, Missoula, Montana)

2. Burgemiester, R. B, Blum, L. H., and Lorge, I. Columbia Mental Maturity Scale, 1959 Revision. World Book Company: Yonkers-on-Hudson, 1959.

3. Dunn, L. M. Peabody Picture Vocabulary Test Manual. Tentative edition. Nashville, Tennessee: George Peabody College for Teachers, 1958.

4. Van Alystyne, D. Van Alystyne Picture Vocabulary Test, 1959 Revision, World Book Company: Yonkers-on-Hudson, 1959.

PEABODY PICTURE VOCABULARY TEST

Individual Test Record

Name ______________________________ Sex: M F Grade ________
(Last) (First) (Initial) (circle)

School ______________________ Teacher ______________________
(or address) (or parent or phone)

Calculation		DERIVED SCORES			Year	Month	Day
Ceiling item	______	Age equivalent (V.A.)	______	Date	______	______	______
Errors	______	Standard score (V.Q.)	______	Born	______	______	______
Raw score	______	Percentile (%ile)	______	Age	______	______	

Examiner ______________________ Time ____________ Code ____________

• • • J-1 • • • F-2 • • • M-3 • • • A-4 • • • M-5 • • • J-6 • • • J-7 • • • A-8 • • • S-9 • • • O-10 • • • N-11 • • • D-12 • • •

TEST BEHAVIOR

Examples needed:	____only 1	____2 or 3	____over 3
Type of response:	____Subject pointed	____S. called numbers	____Examiner pointed
Rapport:	____easily attained	____slowly attained	____poor rapport
Guessing:	____prone to guess	____guessed when asked	____resisted guessing
Speed of response:	____fast	____average	____slow
Verbalization:	____talkative	____average	____taciturn
Attention span:	____distractible	____average	____very attentive
Perseveration:	____none noted	____some	____frequent
Need for praise:	____little needed	____some needed	____much needed

Other test behavior:

PHYSICAL CHARACTERISTICS

Motor activity:	____hyperactive	____average	____hypoactive
Sedation:	____none	____slight	____heavy
Ambulation:	____normal	____walks with support	____none
Speech:	____intelligible	____fairly intelligible	____unintelligible
Hearing: necessity to repeat stimulus words	____never	____seldom	____often
	____S. wore hearing aid	____S. watched examiner's lips and face closely.	
Vision: distance of eyes from page	____under 8"	____average (8"-20")	____over 20"
	____S. wore glasses	____S. owned but did not wear glasses during test.	

Other physical characteristics:

OTHER INFORMATION (previous tests, dates, scores etc.; teacher estimates of vocabulary, intelligence, achievement; school or work record)

All rights reserved. This form may not be reproduced.

Name____

Form B

Item	Resp.	Key Word
1		(2) table
2		(4) bus
3		(2) horse
4		(3) dog
5		(4) shoe
6		(4) finger
7		(3) boat
8		(2) children
9		(1) bell
10		(4) turtle
11		(2) climbing
12		(1) lamp
13		(3) sitting
14		(2) jacket
15		(1) pulling
16		(2) ring
17		(1) nail
18		(2) hitting
19		(3) tire
20		(3) ladder
21		(1) snake
22		(1) river
23		(4) ringing
24		(4) baking
25		(2) cone
26		(3) engineer
27		(4) peeking
28		(1) kite
29		(1) rat
30		(1) time
31		(4) sail
32		(2) ambulance
33		(2) trunk
34		(4) skiing
35		(2) hook
36		(1) tweezers
37		(3) wasp
38		(2) barber
39		(3) parachute
40		(4) saddle
41		(3) temperature
42		(1) captain
43		(2) whale
44		(4) cash
45		(1) balancing
46		(3) cobweb
47		(3) pledging
48		(1) argument
49		(3) hydrant
50		(4) binocular
51		(1) locomotive
52		(2) hive
53		(4) reel
54		(1) insect
55		(1) gnawing
56		(2) weapon
57		(3) bannister
58		(1) idol
59		(1) globe
60		(3) walrus
61		(1) filing
62		(3) shears
63		(1) horror
64		(4) chef
65		(4) harvesting
66		(3) construction
67		(4) observatory
68		(4) assistance
69		(2) erecting
70		(3) thoroughbred
71		(2) casserole
72		(4) ornament
73		(3) cobbler
74		(2) autumn
75		(3) dissatisfaction
76		(4) scholar
77		(1) oasis
78		(3) soldering
79		(3) astonishment
80		(1) tread
81		(2) thatched
82		(1) jurisprudence
83		(2) sapling
84		(3) arch
85		(4) dwelling
86		(1) lubricating
87		(2) pedestrian
88		(3) vale
89		(3) jubilant
90		(2) laden
91		(2) pursuit
92		(4) goblet
93		(2) rodent
94		(3) confiding
95		(4) reclining
96		(1) frisking
97		(2) moat
98		(3) salutation
99		(2) barrier
100		(3) foal
101		(4) incandescent
102		(3) cornucopia
103		(2) ascending
104		(1) summit
105		(3) caster
106		(2) lobe
107		(3) patriarch
108		(3) sampler
109		(3) ingenious
110		(1) repose
111		(3) constrain
112		(1) tangent
113		(4) sconce
114		(4) hoary
115		(1) pendant
116		(1) prodigy
117		(2) casement
118		(1) quiescent
119		(4) talon
120		(1) chevron
121		(4) feline
122		(2) cairn
123		(4) convergence
124		(3) apothecary
125		(2) indigent
126		(4) edifice
127		(3) scallion
128		(1) infirm
129		(1) emaciate
130		(2) catapult
131		(2) arable
132		(4) orifice
133		(3) renovate
134		(1) precarious
135		(2) dromedary
136		(1) pedagogue
137		(1) sepal
138		(3) lethargic
139		(4) delectation
140		(3) embellish
141		(1) osculation
142		(2) cincture
143		(3) barrister
144		(3) carrion
145		(2) lanate
146		(4) chirography
147		(1) mendicant
148		(1) saltation
149		(2) florescence
150		(4) culver

PEABODY PICTURE VOCABULARY TEST

Individual Test Record

Name ____________________ (Last) ____________________ (First) ____________________ (Initial) Sex: M F (circle) Grade __________

School ______________________________ (or address) Teacher ______________________________ (or parent or phone)

Calculation		DERIVED SCORES			Year	Month	Day
Ceiling item	______	Age equivalent (V.A.)	______	Date	______	______	______
Errors	______	Standard score (V.Q.)	______	Born	______	______	______
Raw score	______	Percentile (%ile)	______	Age	______	______	

Examiner ______________________________ Time ______________ Code ______________

• • • J-1 • • • F-2 • • • M-3 • • • A-4 • • • M-5 • • • J-6 • • • J-7 • • • A-8 • • • S-9 • • • O-10 • • • N-11 • • • D-12 • • •

TEST BEHAVIOR

Examples needed:	____only 1	____2 or 3	____over 3
Type of response:	____Subject pointed	____S. called numbers	____Examiner pointed
Rapport:	____easily attained	____slowly attained	____poor rapport
Guessing:	____prone to guess	____guessed when asked	____resisted guessing
Speed of response:	____fast	____average	____slow
Verbalization:	____talkative	____average	____taciturn
Attention span:	____distractible	____average	____very attentive
Perseveration:	____none noted	____some	____frequent
Need for praise:	____little needed	____some needed	____much needed

Other test behavior:

PHYSICAL CHARACTERISTICS

Motor activity:	____hyperactive	____average	____hypoactive
Sedation:	____none	____slight	____heavy
Ambulation:	____normal	____walks with support	____none
Speech:	____intelligible	____fairly intelligible	____unintelligible
Hearing: necessity to repeat stimulus words	____never	____seldom	____often
	____S. wore hearing aid	____S. watched examiner's lips and face closely.	
Vision: distance of eyes from page	____under 8"	____average (8"-20")	____over 20"
	____S. wore glasses	____S. owned but did not wear glasses during test.	

Other physical characteristics:

OTHER INFORMATION

(previous tests, dates, scores etc.; teacher estimates of vocabulary, intelligence, achievement; school or work record)

All rights reserved. This form may not be reproduced.

Name____ **Form A**

Item	Resp.	Key Word
1		(4) car
2		(3) cow
3		(1) baby
4		(2) girl
5		(1) ball
6		(3) block
7		(2) clown
8		(1) key
9		(4) can
10		(2) chicken
11		(4) blowing
12		(2) fan
13		(1) digging
14		(1) skirt
15		(4) catching
16		(1) drum
17		(3) leaf
18		(4) tying
19		(1) fence
20		(2) bat
21		(4) bee
22		(3) bush
23		(1) pouring
24		(1) sewing
25		(4) wiener
26		(2) teacher
27		(3) building
28		(3) arrow
29		(2) kangaroo
30		(3) accident
31		(3) nest
32		(4) caboose
33		(1) envelope
34		(2) picking
35		(1) badge
36		(3) goggles
37		(2) peacock
38		(3) queen
39		(4) coach
40		(1) whip
41		(4) net
42		(4) freckle
43		(3) eagle
44		(2) twist
45		(4) shining
46		(2) dial
47		(2) yawning
48		(2) tumble
49		(1) signal
50		(1) capsule
51		(4) submarine
52		(4) thermos
53		(3) projector
54		(4) group
55		(3) tackling
56		(1) transportation
57		(1) counter
58		(2) ceremony
59		(3) pod
60		(4) bronco
61		(3) directing
62		(4) funnel
63		(2) delight
64		(3) lecturer
65		(2) communication
66		(4) archer
67		(1) stadium
68		(1) excavate
69		(4) assaulting
70		(1) stunt
71		(1) meringue
72		(3) appliance
73		(4) chemist
74		(3) arctic
75		(4) destruction
76		(3) porter
77		(2) coast
78		(4) hoisting
79		(1) wailing
80		(2) coil
81		(3) kayak
82		(2) sentry
83		(4) furrow
84		(1) beam
85		(3) fragment
86		(2) hovering
87		(3) bereavement
88		(4) crag
89		(2) tantrum
90		(1) submerge
91		(3) descend
92		(2) hassock
93		(1) canine
94		(1) probing
95		(1) angling
96		(3) appraising
97		(4) confining
98		(4) precipitation
99		(1) gable
100		(1) amphibian
101		(3) graduated
102		(2) hieroglyphic
103		(1) orate
104		(3) cascade
105		(4) illumination
106		(1) nape
107		(2) geneologist
108		(2) embossed
109		(4) mercantile
110		(2) encumbered
111		(4) entice
112		(3) concentric
113		(3) vitreous
114		(1) sibling
115		(2) machete
116		(4) waif
117		(1) cornice
118		(3) timorous
119		(1) fettered
120		(2) tartan
121		(3) sulky
122		(4) obelisk
123		(2) ellipse
124		(2) entomology
125		(4) bumptious
126		(2) dormer
127		(2) coniferous
128		(4) consternation
129		(3) obese
130		(4) gauntlet
131		(1) inclement
132		(1) cupola
133		(2) obliterate
134		(3) burnishing
135		(1) bovine
136		(4) eminence
137		(3) legume
138		(4) senile
139		(2) deleterious
140		(4) raze
141		(2) ambulation
142		(1) cravat
143		(2) impale
144		(4) marsupial
145		(3) predatory
146		(1) incertitude
147		(2) imbibe
148		(3) homunculus
149		(4) cryptogam
150		(3) pensile